



TESIS DOCTORAL

MECANISMOS PARA LA GESTIÓN EFICIENTE DEL PLANO DE CONTROL Y DEL PLANO DE DATOS EN REDES MÓVILES 5G

JESÚS MANUEL CALLE CANCHO

PROGRAMA DE DOCTORADO EN TECNOLOGÍAS INFORMÁTICAS

2020



TESIS DOCTORAL

MECANISMOS PARA LA GESTIÓN EFICIENTE DEL PLANO DE CONTROL Y DEL PLANO DE DATOS EN REDES MÓVILES 5G

Jesús Manuel Calle Cancho

Conformidad de los Directores:

La conformidad de los Directores de la tesis consta en
el original en papel de esta Tesis Doctoral

José Luis González Sánchez

Javier Carmona Murillo

PROGRAMA DE DOCTORADO EN TECNOLOGÍAS INFORMÁTICAS

2020

Agradecimientos

En primer lugar, me gustaría mostrar mi agradecimiento a mis directores, José Luis González Sánchez y Javier Carmona Murillo, por su apoyo y ayuda a lo largo de este periodo.

Además, agradecer a mis compañeros del grupo de investigación GÍTACA y del centro CénitS, porque todos me han enseñado algo; especialmente a David Cortés, por sus consejos y apoyo durante el desarrollo de la Tesis.

Por último, mi agradecimiento más especial a mi familia. A mi madre, por todo y por tanto, te estaré eternamente agradecido por darme todo lo necesario para ser feliz, por apoyarme y confiar en mí hasta el último momento. Porque tú lo querías así, mamá. A mi padre, por aceptar y apoyar todas y cada una de mis decisiones, enseñándome que con esfuerzo, trabajo y humildad todo se consigue. A María José, por su apoyo incondicional, su paciencia y su cariño. Porque contigo todo es más fácil, las alegrías a tu lado se doblan y la tristeza se reduce a la mitad. Solamente tú sabes lo que me ha costado llegar hasta aquí: *Always*.

Resumen

En los últimos años, el incremento exponencial del tráfico de datos móviles, unido al despliegue de nuevos servicios sobre las redes actuales, han propiciado que los operadores de telecomunicaciones busquen nuevos mecanismos que permitan una gestión eficiente de la red de acuerdo a las demandas específicas de los usuarios. Estos mecanismos deben ser capaces de gestionar los recursos de red dinámicamente, proporcionando flexibilidad en el nuevo entorno 5G, a través de tecnologías emergentes. En este contexto, uno de los procesos involucrados en la gestión de la red es el soporte a la movilidad, cuyo principal objetivo es mantener las comunicaciones activas mientras los usuarios se mueven entre redes de acceso diferentes de manera transparente, minimizando el tiempo de desconexión o pérdida de servicio.

A tal efecto, se han estandarizado protocolos para la gestión de la movilidad centralizada (CMM) que basan su funcionamiento en una entidad central que gestiona el tráfico de datos y la señalización de los nodos móviles. Sin embargo, estos protocolos presentan algunas limitaciones y necesitan ser adaptados a las nuevas tendencias y a la creciente demanda de tráfico de datos móviles. Por ello, se han desarrollado nuevas soluciones de gestión de la movilidad distribuida (DMM), cuyo objetivo es distribuir los nodos que actúan como ancla de movilidad por el borde de la red más cerca de los usuarios finales.

Por lo tanto, DMM reduce algunos de los problemas que afectan a las redes móviles, pero debido a la densificación de celdas producida por el incipiente desarrollo de 5G, se está produciendo un incremento de tráfico de señalización usado para gestionar la movilidad, que debe ser tenido en cuenta por los operadores de red en la fase de diseño. Además, a pesar de que muchos protocolos se están desarrollando de forma distribuida, existen situaciones en las que DMM provoca mayores costes de la red y, por tanto, su rendimiento puede verse afectado.

Partiendo de esta situación, en esta tesis se proponen tres nuevos mecanismos para mejorar el rendimiento de las redes móviles de próxima generación desde tres

perspectivas diferentes. Nuestra primera propuesta, TE-DMM, permite llevar a cabo una gestión eficiente del plano de control, reduciendo en gran medida el tráfico de señalización. La segunda propuesta, SR-DMM, combina SDN con DMM para mejorar el proceso de gestión de la movilidad desde el punto de vista del plano de datos. Nuestro tercer mecanismo propone una estrategia de asociación entre estaciones base y la red de acceso para mejorar el rendimiento de los protocolos de gestión de la movilidad, tanto del plano de control como del plano de datos.

Finalmente, se ha diseñado e implementado un simulador de gestión de la movilidad para redes de nueva generación (PyMMSim), que ha sido utilizado para obtener los resultados de simulación presentados en esta Tesis. Además, se han realizado evaluaciones analíticas y experimentales con el objetivo de medir el rendimiento de los mecanismos propuestos en términos de costes de movilidad.

Keywords— CMM, DMM, protocolos de movilidad, redes móviles de próxima generación, simulación de red, evaluación experimental, costes de movilidad, plano de control, plano de datos.

Lista de Abreviaturas

- 3G** 3th Generation Mobile Networks. 1, 20, 21
- 3GPP** 3rd Generation Partnership Project. 2, 21, 29, 88, 90–92, 121, 152
- 4G** 4th Generation Mobile Networks. 1, 20, 21, 29
- 5G** 5th Generation Mobile Networks. III, 1, 20, 21, 25, 29, 37, 67, 68, 84, 88–90, 109, 129, 152, 154–156
- 5G PPP** 5G Infrastructure Public Private Partnership. 24, 30, 89, 156
- 5GC** 5G Core Network. 29
- ABA** Access Binding Acknowledgement. 34
- ABU** Access Binding Update. 34
- AMA** Access Mobility Anchor. 34
- API** Interfaz de Programación de Aplicaciones. 28
- AR** Access Router. 31, 91
- AT&T** American Telephone and Telegraph. 27
- ATM** Asynchronous Transfer Mode. 27
- BA** Binding Acknowledgement. 31
- BS** Base Station. 114, 117, 118, 120, 125
- BU** Binding Update. 31, 34
- C-RAN** Cloud Radio Access Network. XIII, 117, 119, 120, 125–128, 154

CAPEX Capital Expenditures. 13, 23, 24, 111, 113–115, 117, 119, 122, 124, 125, 127, 129, 148, 154, 155

CDF Función de Distribución Acumulada. 63, 82

CMD Context Mobility Database. 36, 45–47, 52–54, 60–62, 69, 80, 81, 93

CMM Centralized Mobility Management. III, 25, 30, 33, 89, 99, 100, 104, 105, 107

CN Correspondent Node. 33, 50, 61, 62, 73, 76, 80–82, 100, 148

CoA Care-of Address. 31

CUPS Control and User Plane Separation. 29

DMM Distributed Mobility Management. III, IV, 13, 25, 29, 33, 34, 36, 37, 45, 89, 92, 93, 99–101, 104, 105, 107, 153

DMM-GW DMM gateway. 92, 93, 100, 101

eMBB enhanced Mobile Broadband. 21

EPC Evolved Packet Core. 29, 112, 121–124, 128, 129

ETSI European Telecommunications Standards Institute. 24

ForCES Forwarding and Control Element Separation. 27

HA Home Agent. 30, 31, 33, 34

HB-DMM Host-Based DMM. 34, 35

HCPP Hard Core Point Process. 139, 142

HN Home Network. 30, 31

HNP Home Network Prefix. 35, 93

HoA Home Address. 30

HPC High Performance Computing. 144, 145

HSS Home Subscriber Server. 121

IETF Internet Engineering Task Force. 2, 4, 20, 27, 29, 30, 152

IoT Internet of Things. 2

IRSCP Intelligent Route Service Control Point. 27

ISG Industry Specification Group. 24

LMA Local Mobility Anchor. 31–33, 91

LNA Link-Network Assignment. 88, 95, 98, 101–105, 107, 109, 137, 140, 153

MAD-PMIPv6 Mobility Anchors Distribution for Proxy Mobile IPv6. 58, 79

MAG Mobile Access Gateway. 31, 32

MAR Mobility Access Router. 35, 36, 45, 46, 48–50, 52, 53, 55, 57, 61–63, 70, 75, 80–83

MEC Mobile Edge Computing. 1, 21, 24, 89, 156

MIoT Massive Internet of Things. 21

MIPv6 Mobile IPv6. 30, 31, 33, 34, 39, 141

MME Mobility Management Entity. 121

MN Mobile Node. XII, 30–36, 38, 39, 45–50, 52, 54–57, 59, 60, 62–64, 69–76, 78, 80–85, 91–93, 95, 99, 100, 141, 146, 148

NB-DMM Network-Based DMM. 35, 36, 44–46, 48–50, 52, 54, 56–58, 61, 62, 64, 66, 69, 70, 73–85, 90, 92, 99, 141, 152, 153

NFV Network Function Virtualization. 1, 13, 21, 23, 24, 29, 111–114, 121, 128, 129, 154, 156

OPEX Operational Expenditures. 13, 23, 24, 111, 113–115, 118, 120, 122–124, 126, 127, 129, 148, 154, 155

PBA Proxy Binding Acknowledgement. 31, 36, 46, 47, 59, 60, 74, 76, 91, 93

PBU Proxy Binding Update. 31, 36, 46–48, 60, 74, 76, 91, 93

PCE Path Computation Element. 27

PDN Packet Data Network. 121

PGW Packet Data Network Gateway. 91, 92, 100, 121, 155

PMIPv6 Proxy Mobile IPv6. 30–33, 35, 90–92, 99, 141

PP Point Process. 138

PPP Poisson Point Process. 102, 138

PyMMSim Python Mobility Management Simulator. IV, 131, 132, 135–149, 154–156

QoE Quality of Experience. 21, 25, 33, 89, 155

QoS Quality of Service. 21, 25, 33, 89, 155

RA Router Advertisement. 46, 47, 60, 70, 71, 73, 77, 83

RAN Radio Access Networks. 117, 118, 120, 125, 126, 128

RCP Routing Control Platform. 27

RCT Red Científico Tecnológica. 112, 123, 124

RQ Research Question. 10, 11

RRH Remote Radio Head. 117

RS Router Solicitation. 38, 46, 47, 70, 71, 77, 83

SBS Software Defined Base Station. 114, 117, 119, 120, 125, 128

SDN Software Defined Networking. IV, 1, 11–13, 21–23, 26, 28, 29, 37–39, 65, 67–71, 73, 75, 78–80, 83, 84, 111–114, 128, 129, 153, 154

SFC Service Function Chaining. 154

SGW Serving Gateway. 91, 92, 100, 121

SLR Systematic Literature Review. 4–6, 11

SR-DMM Software Redirection to Distributed Mobility Management. IV, XII, 65, 66, 68–85, 153

T-MAR Tunneling-Mobility Access Router. 45, 47, 48, 53, 55–63

TCO Total Cost Ownership. 124, 127, 129, 154

TCP Transmission Control Protocol. 78

TE-DMM Tunneling Extension to Distributed Mobility Management. IV, XII, 13, 43–45, 47–50, 53, 54, 56–62, 64, 152, 153

URLLC Ultra-Reliable and Low-Latency Communication. 21

VBS Virtualized Base Station. 117, 119, 120, 125

VNF Virtual Network Function. 23

VPN Virtual Private Networks. 23

Índice de Contenidos

Resumen	III
Lista de Abreviaturas	V
Lista de Figuras	XV
Lista de Tablas	XIX
1. Introducción	1
1.1. Planteamiento del problema	1
1.2. Objetivos	3
1.3. Revisión sistemática de la literatura	4
1.3.1. Fase I: Planificación	5
1.3.2. Fase II: Ejecución	7
1.3.3. Fase III: Resultados	8
1.4. Estructura de la Tesis	10
1.5. Contribuciones	12
1.6. Publicaciones	14
1.7. Participación en proyectos de investigación	16
2. Redes Móviles de Próxima Generación	19
2.1. Introducción	20
2.2. Evolución de las arquitecturas de red	26

2.2.1.	Primeros esfuerzos de separación de los planos de control y de datos	26
2.2.2.	Redes programables: OpenFlow	28
2.2.3.	Redes de próxima generación y gestión de la movilidad	29
2.3.	Soporte a la gestión de la movilidad	29
2.3.1.	Gestión de la movilidad centralizada	30
2.3.2.	Gestión de la movilidad distribuida	32
2.3.3.	Retos de la gestión de la movilidad	37
2.3.4.	Comparativa entre protocolos de gestión de la movilidad	38
2.3.5.	Principales métricas para el análisis de los protocolos de gestión de la movilidad	39
3.	Propuesta para la Mejora del Plano de Control en Redes Móviles	43
3.1.	Arquitectura propuesta TE-DMM	45
3.2.	Modelo analítico	50
3.2.1.	Coste de señalización	52
3.2.2.	Coste de procesamiento	53
3.2.3.	Pérdida de paquetes	54
3.3.	Evaluación de rendimiento	54
3.4.	Evaluación experimental	58
3.5.	Conclusiones	64
4.	Propuesta para la Mejora del Plano de Datos en Redes Móviles	65
4.1.	Enfoque de gestión de movilidad basado en SDN	67
4.1.1.	Solución SR-DMM	68
4.1.2.	Registro de un MN con SR-DMM	70
4.1.3.	Proceso de <i>handover</i> con SR-DMM	71
4.2.	Modelo analítico	73
4.2.1.	Coste de señalización	74

4.2.2. Coste de entrega de los paquetes de datos	76
4.3. Resultados obtenidos	77
4.4. Evaluación experimental	79
4.5. Conclusiones	84
5. Mecanismo de asignación Link-Network para la Mejora de los Planos de Control y de Datos	87
5.1. Introducción	88
5.2. Gestión de la movilidad en arquitecturas 3GPP	90
5.2.1. Gestión de la movilidad centralizada en redes de acceso 3GPP .	90
5.2.2. Gestión de la movilidad distribuida en redes de acceso 3GPP . .	92
5.3. Modelado del sistema y formulación del problema	93
5.4. Métricas de evaluación	99
5.4.1. Plano de control	99
5.4.2. Plano de datos	100
5.5. Evaluación de rendimiento	101
5.6. Conclusiones	109
6. Análisis de Costes del Despliegue de una Arquitectura de Red Virtualizada	111
6.1. Introducción	112
6.2. Virtualización de red	113
6.3. Modelo de costes	115
6.3.1. CAPEX y OPEX en una red de acceso tradicional	117
6.3.2. CAPEX y OPEX en una red de acceso virtualizada (C-RAN) . .	119
6.3.3. CAPEX y OPEX en una red de acceso 3GPP virtualizada	121
6.4. Caso RCT: resultados obtenidos	123
6.5. Conclusiones	128

7. PyMMSim: Simulador de Gestión de la Movilidad en Redes de Nueva Generación	131
7.1. Introducción	132
7.2. Estructura de PyMMSim	135
7.3. Escenario de simulación	142
7.4. Entorno de ejecución	144
7.5. Métricas para evaluar la gestión de la movilidad	147
7.6. Conclusiones	149
8. Conclusiones y Trabajo Futuro	151
8.1. Conclusiones	151
8.2. Trabajo futuro	154
A. Infraestructura Utilizada en las Simulaciones	157
Referencias	159

Lista de Figuras

1.1. Líneas de actuación de la Tesis	3
1.2. Proceso de revisión sistemática de la literatura	5
1.3. Número de artículos incluidos durante el proceso de selección	8
1.4. Número de artículos por fuente	9
1.5. Número de artículos aceptados y rechazados por fuente	9
1.6. Número total de artículos por año	10
1.7. Organización de la Tesis en base a las preguntas de investigación	10
1.8. Organización de la Tesis y relación entre los capítulos	11
2.1. Comparativa entre una arquitectura de red tradicional y una arquitectura de red SDN	22
2.2. Arquitectura de red de NFV	24
2.3. Evolución de la separación entre el plano de datos y el plano de control	26
2.4. Visión general de la arquitectura de MIPv6	31
2.5. Visión general de la arquitectura de PMIPv6	32
2.6. Visión general de la arquitectura de HB-DMM	35
2.7. Visión general de la arquitectura de NB-DMM parcialmente distribuida	36
2.8. Visión general de la arquitectura de NB-DMM totalmente distribuida .	37
2.9. Visión general de la arquitectura de SDN-DMM	38
3.1. Arquitectura de NB-DMM	44
3.2. Visión general de la arquitectura de red propuesta: TE-DMM	45

3.3. Diagrama de flujo de mensajes en NB-DMM	46
3.4. Diagrama de flujo de mensajes en TE-DMM	47
3.5. Primera fase de gestión de flujos: conexión del MN al MAR1/T-MAR1 .	48
3.6. Segunda fase de gestión de flujos: movimiento del MN al MAR2/T-MAR2	48
3.7. Tercera fase de gestión de flujos: movimiento del MN al MAR3/T-MAR3. Creación de nuevos túneles (a) y extensión de túnel de T-MAR2 a T-MAR3 (b).	49
3.8. Cuarta fase de gestión de flujos: movimiento del MN al MAR4/T-MAR4. Creación de nuevos túneles (a) y extensión de túnel de T-MAR3 a T-MAR4 (b).	49
3.9. Distribución de celdas en un área de cobertura	51
3.10. Coste de señalización dependiendo de la velocidad del MN	55
3.11. Coste de señalización dependiendo de la longitud del lado de la celda hexagonal	56
3.12. Coste de procesamiento dependiendo del número de anclas de movilidad	57
3.13. Pérdida de paquetes dependiendo del número medio de paquetes de una sesión	58
3.14. Procedimiento de extensión de túneles en la propuesta TE-DMM	59
3.15. <i>Testbed</i> usado para la evaluación experimental de TE-DMM y NB-DMM	61
3.16. Función de distribución acumulada (CDF) experimental de las medidas de latencia <i>handover</i>	63
3.17. Evaluación experimental de paquetes perdidos	63
4.1. Arquitectura de NB-DMM	66
4.2. Arquitectura funcional de SDN	67
4.3. Arquitectura funcional de la propuesta SR-DMM	69
4.4. Diagrama de flujo de mensajes: registro inicial en SR-DMM	70
4.5. Diagrama de flujo de mensajes: proceso de <i>handover</i> en SR-DMM . . .	71
4.6. Topología utilizada para el modelo analítico de NB-DMM y SR-DMM .	74
4.7. Coste de señalización con respecto al tiempo de estancia en una celda .	78

4.8. Coste de entrega de paquetes con respecto al tiempo de estancia en una celda	79
4.9. Topología del <i>testbed</i> utilizado en la evaluación experimental	80
4.10. Función de distribución acumulada (CDF) experimental de las medidas de latencia <i>handover</i>	82
4.11. Latencia de <i>handover</i> segregada para NB-DMM y SR-DMM	83
5.1. Visión general de la arquitectura de PMIPv6 sobre la red de acceso 3GPP	91
5.2. Visión general de la arquitectura de NB-DMM sobre la red de acceso 3GPP	92
5.3. Asignación de las estaciones base al dominio de movilidad	94
5.4. Ejemplo del algoritmo propuesto LNA para tres <i>routers</i> de acceso . . .	98
5.5. Evaluación de complejidad de computación del algoritmo propuesto LNA	99
5.6. Topología de red utilizada en las simulaciones con ocho <i>routers</i> de acceso ($K = 8$)	102
5.7. Evaluación del plano de control en protocolos CMM usando diferentes algoritmos de asignación	104
5.8. Evaluación del plano de control en protocolos DMM usando diferentes algoritmos de asignación	105
5.9. Evaluación del plano de datos en protocolos CMM usando diferentes algoritmos de asignación	106
5.10. Evaluación del plano de datos en protocolos DMM usando diferentes algoritmos de asignación	106
6.1. Relación entre SDN y NFV	114
6.2. Costes de infraestructura y operación considerados	116
6.3. Comparativa entre una RAN tradicional y una RAN virtualizada (C-RAN)	116
6.4. Red Científico Tecnológica virtualizada	124
6.5. Análisis de CAPEX: RAN vs C-RAN	126
6.6. Análisis de OPEX: RAN vs C-RAN	126
6.7. Análisis de TCO: RAN vs C-RAN	127

6.8. Comparativa de costes relativos de RAN vs C-RAN en función de la densidad de usuarios	127
6.9. Análisis del CAPEX y OPEX del EPC virtualizado	128
7.1. Maneras de estudiar un sistema	133
7.2. Principio de la simulación de eventos discretos. Durante la simulación el estado del sistema cambia solo en puntos discretos de tiempo t_i	134
7.3. Diagrama de flujo del algoritmo genérico de tratamiento de eventos discretos	134
7.4. Diseño modular del simulador PyMMSim	135
7.5. Simulación de eventos discretos con PyMMSim	136
7.6. Procesos de puntos implementados en PyMMSim	138
7.7. Área de cobertura de las estaciones base en PyMMSim según una teselación de Voronoi	139
7.8. Distribución de las demandas en el tiempo de simulación	141
7.9. Topologías de red implementadas en PyMMSim	142
7.10. Integración de PyMMSim en un entorno HPC	144
7.11. Estructura del directorio de resultados de PyMMSim	146

Lista de Tablas

1.1. <i>Búsqueda en bases de datos</i>	7
1.2. <i>Estudios importados de cada librería digital</i>	8
2.1. <i>Equivalencia entre los principales roles de movilidad y entidades lógicas</i>	32
2.2. <i>Comparativa entre protocolos de gestión de movilidad centralizada y distribuida</i>	39
2.3. <i>Resumen de trabajos de análisis y evaluación de gestión de la movilidad</i>	41
4.1. <i>Notación utilizada en el modelo analítico</i>	76
4.2. <i>Media y desviación típica de los resultados obtenidos de latencia de handover</i>	82
5.1. <i>Funciones de los nodos de la topología de red para $K = 8$</i>	102
5.2. <i>Evaluación de los protocolos de gestión de la movilidad usando diferentes algoritmos de asignación entre las estaciones base y los routers de acceso.</i>	108
6.1. <i>Parámetros tenidos en cuenta para la evaluación</i>	125
7.1. <i>Métricas para la evaluación de la gestión de la movilidad en PyMMSim</i>	148

Capítulo 1

Introducción

Este capítulo proporciona una visión general y antecedentes de la investigación llevada a cabo en esta Tesis, detallando los objetivos que se pretenden conseguir, las contribuciones realizadas y la estructura de la misma.

1.1. Planteamiento del problema

En los últimos años, el principal objetivo de los proveedores de servicios de telecomunicación ha sido dar cobertura a la creciente demanda llevada a cabo por sus usuarios [1]. Debido a ello, ha surgido una nueva generación de redes de comunicaciones móviles (5G) [2], que no solo se plantea como una evolución de las redes 3G y 4G, sino que soporta una amplia variedad de servicios y que aporta mejoras desde el punto de la banda ancha móvil, proporcionando alta fiabilidad y baja latencia en las comunicaciones. De esta manera, 5G dará respuesta a retos más complejos, lo que exige un avance tecnológico significativo con respecto al estado actual de la tecnología [3]. En definitiva, se espera que estas redes móviles de próxima generación sean capaces de proporcionar nuevos servicios de acuerdo a las demandas específicas de los usuarios.

Para hacer frente a estos desafíos, tanto la industria como la academia han impulsado el funcionamiento de la red basado en mecanismos emergentes como las redes definidas por *software* (SDN, *Software Defined Networks*) o la virtualización de las funciones de red (NFV, *Network Function Virtualization*), con los que se espera alcanzar el rendimiento necesario y la escalabilidad y flexibilidad esperada [4]. Además, las funciones se ejecutarán en un sistema operativo unificado, especialmente en los bordes de la red bajo el paradigma de *Mobile Edge Computing* (MEC) [5]. En definitiva,

todos ellos son intentos continuos para enfrentarse a los retos de la nueva era de redes de datos móviles.

Por lo tanto, estos cambios implican adaptar las redes de acceso y los protocolos de comunicaciones utilizados hasta el momento, para llevar a cabo una gestión eficiente de las redes móviles de próxima generación. Dentro de estas adaptaciones y actuaciones, hay que tener en cuenta el proceso de soporte a la movilidad, que engloba el mantenimiento de la conectividad entre diferentes redes de acceso y la gestión de los recursos requeridos por los usuarios móviles. Para este propósito, las organizaciones más relevantes en el campo de las comunicaciones móviles, como son *Internet Engineering Task Force* (IETF) [6] y *3rd Generation Partnership Project* (3GPP) [7], han diseñado diferentes soluciones de gestión de la movilidad que puedan ser utilizadas sobre arquitecturas de redes móviles. Los paradigmas propuestos más recientemente han surgido por la necesidad que aparece en las redes móviles de próxima generación de reducir la sobrecarga existente en el núcleo de la red [8]. La idea principal es introducir lo que se conoce como arquitecturas de red planas (*flat networks*), en las que las entidades encargadas de gestionar la movilidad se ubiquen más cerca de los usuarios móviles, distribuyendo las funciones de movilidad entre varias entidades de la red de acceso.

Sin embargo, a veces estas soluciones no presentan comportamientos adecuados en ciertos escenarios de movilidad de nueva generación [9] que se ven afectados por el crecimiento del tráfico de datos móviles, que ha aumentado exponencialmente con la aparición del paradigma de *Internet of Things* (IoT) [10, 11].

En este contexto, esta Tesis proporciona una serie de mecanismos con el objetivo de mejorar y reducir los problemas mencionados anteriormente, prestando especial atención a la mejora de rendimiento del plano de control y del plano de datos y de la integración de la red de acceso y el nivel físico; todo ello desde la perspectiva de los protocolos de gestión de la movilidad. En la Figura 1.1 se muestra un esquema de las líneas de actuación comentadas anteriormente y de la arquitectura de red sobre la que se sustentará la investigación desarrollada.

Definición del problema: Las redes móviles de próxima generación requieren mecanismos inteligentes de gestión de red que permitan el control eficiente de la misma y la gestión del tráfico generado por todos los dispositivos móviles que se conectan a ella. Por lo tanto, surge la necesidad de diseñar e implementar nuevos mecanismos que proporcionen mejoras en el rendimiento global de la misma.

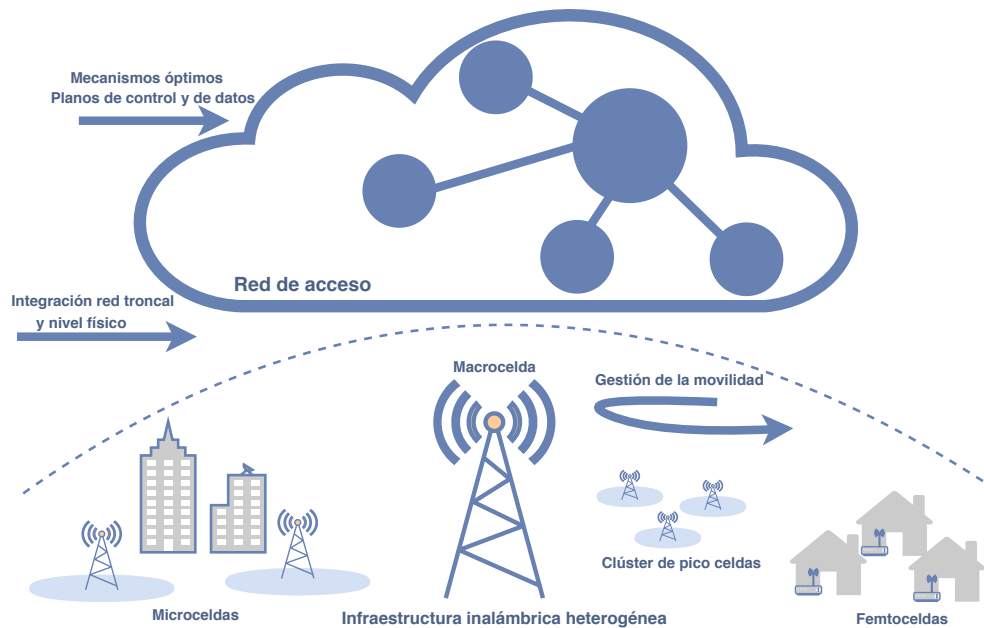


Figura 1.1: Líneas de actuación de la Tesis

1.2. Objetivos

La sección anterior describe las necesidades que han aparecido en las redes móviles de próxima generación debido a los cambios tecnológicos que se han producido en los últimos años. Además, se ha descrito la importancia del soporte a la gestión de la movilidad en estos entornos tan cambiantes. En consecuencia, con el desarrollo de esta Tesis se persigue mejorar la gestión de la movilidad en redes de próxima generación a través del diseño y desarrollo de mecanismos que sean capaces de gestionar la red de un modo óptimo y eficiente. Por lo tanto, el objetivo principal de esta investigación puede ser resumido tal y como se presenta a continuación:

Analizar, diseñar y evaluar mecanismos para la gestión eficiente del plano de control y del plano de datos en redes móviles 5G, para mejorar su acceso en términos de gestión de tráfico, rendimiento y escalabilidad.

Además del objetivo general, se ha detallado una serie de objetivos específicos, tal y como se muestra a continuación:

- Atender y satisfacer los requerimientos y necesidades de los usuarios de las redes móviles de próxima generación, analizando las posibilidades que ofrecen las tecnologías emergentes para mejorar el rendimiento de la red.

- Analizar los mecanismos de gestión de la movilidad en redes heterogéneas de próxima generación desarrollados por el IETF, así como otras soluciones relevantes de la literatura existente.
- Diseñar un mecanismo de gestión de la movilidad distribuida que permita llevar a cabo una gestión eficiente del plano de control.
- Estudiar las oportunidades y principales desafíos relacionados con el paradigma de redes definidas por *software*, para diseñar un mecanismo de gestión de la movilidad distribuida que proporcione mejoras de rendimiento al plano de datos.
- Analizar la integración de la red troncal y el despliegue de nivel físico, diseñando mecanismos óptimos para la mejora de los planos de control y de datos en redes móviles de próxima generación.
- Implementar los mecanismos propuestos en un entorno de simulación de altas prestaciones para evaluar y analizar su rendimiento.

1.3. Revisión sistemática de la literatura

Una de las primeras fases llevada a cabo ha sido una revisión sistemática de la literatura (SLR). Se trata de un método que tiene como objetivo examinar, identificar y evaluar todas las evidencias relacionadas con una pregunta de investigación específica de una forma imparcial [12, 13]. En [14], se define SLR como un método sistemático para identificar, evaluar e interpretar el trabajo de investigadores, académicos y profesionales en un determinado campo.

Por tanto, aparecen diversas motivaciones y razones cuando se pretende abordar una revisión sistemática de la literatura:

- Resumir las evidencias existentes que subyacen en torno a un tema determinado.
- Identificar problemas o limitaciones en la investigación existente para sugerir nuevas áreas de acción.
- Establecer una serie de antecedentes para descubrir nuevas áreas de interés.

El proceso de SLR se lleva a cabo en tres fases que son: planificación, ejecución y resultados. La Figura 1.2 muestra el detalle de las distintas fases que componen el proceso.

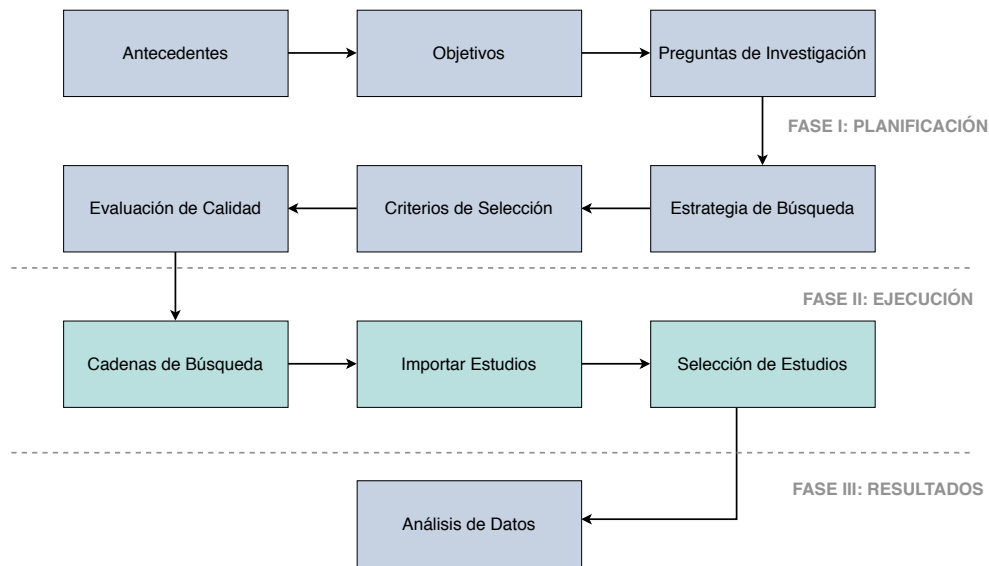


Figura 1.2: Proceso de revisión sistemática de la literatura

1.3.1. Fase I: Planificación

El principal objetivo de esta revisión sistemática de la literatura es identificar trabajos de investigación relacionados con mecanismos para la gestión eficiente del plano de control y del plano de datos en redes móviles de próxima generación, centrándonos en mecanismos de optimización del rendimiento de la red.

Después de definir el objetivo que se pretende conseguir con la SLR, hay que establecer las cuestiones de investigación que la SLR tiene como objetivo responder. Además, se establecerá la estrategia de búsqueda que llevaremos a cabo, así como los criterios de inclusión y exclusión que se aplicarán para realizar uno de los filtrados del proceso.

Preguntas de investigación

Inicialmente es necesario especificar claramente las preguntas de investigación que guiarán esta revisión sistemática de la literatura. Por ello, se han definido las siguientes cuestiones que serán respondidas a través del análisis exhaustivo de los estudios existentes:

1. ¿Cuáles son los retos y oportunidades que han surgido con la evolución hacia las redes móviles de nueva generación?

2. ¿Cuáles son los parámetros críticos de red que necesitan ser optimizados para mejorar el rendimiento de las redes móviles de próxima generación?
3. ¿Qué mecanismos existen para la gestión eficiente de los planos de control y de datos en las redes móviles actuales?

Estrategia de búsqueda

Antes de establecer la cadena de búsqueda se debe definir la estrategia de búsqueda, analizando de manera clara y precisa los términos de búsqueda. Con ello, se identificarán las palabras claves que definen el objetivo principal de la revisión sistemática de la literatura. Por tanto, se han definido las siguientes palabras claves para llevar a cabo la búsqueda de los artículos más relevantes:

- future mobile networks (mobile networks, 4G, 5G).
- mobility management (mobility protocols).
- control plane and data plane.

Si utilizamos los términos anteriores, se puede construir la cadena de búsqueda tal y como se presenta a continuación:

(‘future mobile network’ OR ‘5G’ OR ‘mobile networks’) AND (‘mobility management’ OR ‘mobility protocols’ OR ‘control plane’ OR ‘data plane’)

Con la cadena definida anteriormente, se han llevado a cabo las búsquedas sobre las bases de datos digitales que se muestran a continuación:

- IEEE Digital Library (<https://ieeexplore.ieee.org/>) [15].
- Scopus (<https://www.scopus.com/>) [16].

Criterios de selección

Después de definir los objetivos, las cuestiones de investigación y la estrategia de búsqueda, el siguiente paso en un proceso de SLR consiste en definir los criterios de inclusión y exclusión de los estudios. Estos criterios se usan para excluir los trabajos que no se consideran relevantes para dar respuesta a las preguntas de investigación definidas inicialmente. A continuación se muestran de manera detallada los criterios elegidos:

- Criterios de inclusión:
 - Estudios publicados en revistas.
 - Trabajos publicados entre 2014 y 2020.
- Criterios de exclusión:
 - Estudios no presentados en inglés.
 - Estudios no relacionados con *computer communications and networks*.
 - Estudios relacionados con seguridad.
 - Estudios relacionados con *optical networks*

1.3.2. Fase II: Ejecución

El primer paso dentro de la fase de ejecución ha sido la búsqueda en distintas bases de datos. Cada búsqueda se ha realizado en dos de las fuentes digitales más importantes en la actualidad, como son IEEE Xplore y Scopus. Para cada una de ellas se ha definido una cadena de búsqueda de acuerdo a los criterios específicos que establece cada una, tal y como se muestra en la Tabla 1.1.

Tabla 1.1: Búsqueda en bases de datos

Base de datos	Cadena de búsqueda
IEEE Xplore	('future mobile networks' OR '5G' OR 'mobile networks') AND ('control plane' OR 'data plane' OR 'mobility management' OR 'mobility protocols')
Scopus	TITLE-ABS-KEY (('future mobile networks' OR '5G' OR 'mobile networks') AND ('control plane' OR 'data plane' OR 'mobility management' OR 'mobility protocols'))

Una vez que se ha realizado la búsqueda en cada una de las librerías digitales, se ha llevado a cabo la importación de los estudios desde la herramienta Parsifal [17]; una herramienta diseñada para dar soporte a los investigadores en el proceso de revisión sistemática de la literatura, siendo capaz de gestionar un gran número de referencias bibliográficas de librerías digitales como IEEE Xplore y Scopus.

La Tabla 1.2 muestra el número de estudios importados de cada base de datos.

Tabla 1.2: Estudios importados de cada librería digital

Fuente	Estudios importados
IEEE Xplore	356
Scopus	333

1.3.3. Fase III: Resultados

El principal objetivo de esta fase es detectar artículos relevantes que conformen el resultado final de la revisión sistemática de la literatura y, finalmente, se pueda dar respuesta a las preguntas de investigación formuladas inicialmente.

La Figura 1.3 muestra el proceso seguido durante la selección de los estudios, detallándose el número de artículos incluidos y excluidos.

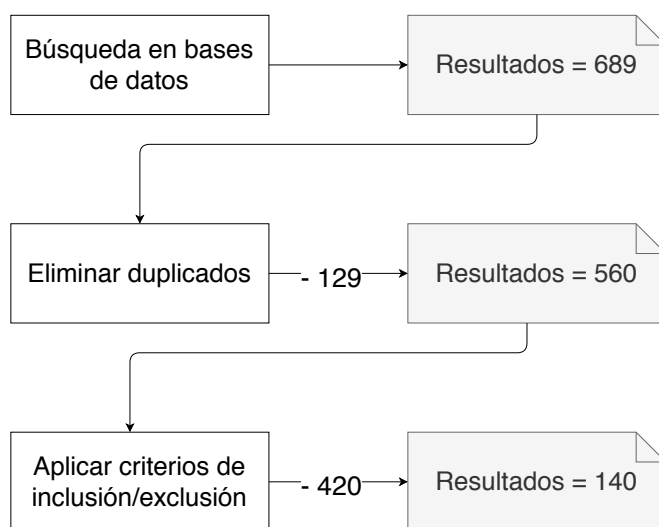


Figura 1.3: Número de artículos incluidos durante el proceso de selección

Después de realizar la búsqueda inicial en las librerías digitales se ha obtenido un total de 689 artículos, de los cuales se han eliminado 129 por estar duplicados. A continuación, se han aplicado los criterios de inclusión y exclusión, resultando un total de 140 artículos seleccionados.

La Figura 1.4 muestra el porcentaje de artículos extraídos de cada fuente después de llevar a cabo la primera búsqueda y eliminar duplicados. El 49,5 % de los 560 artículos pertenecen a IEEE Xplore y el 50,5 % restante a Scopus.

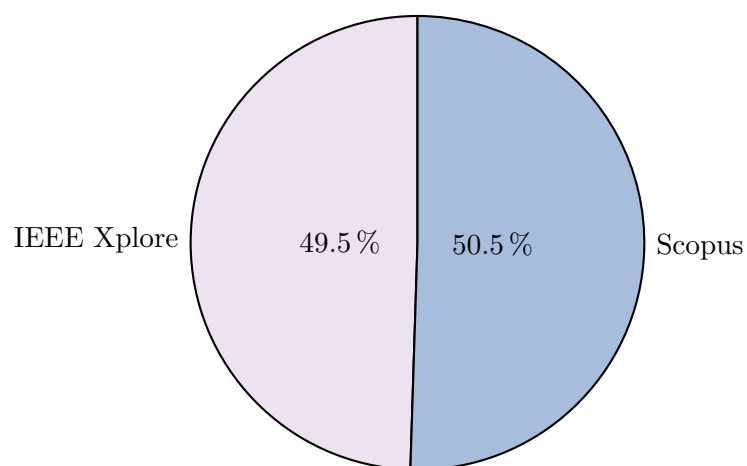


Figura 1.4: Número de artículos por fuente

La Figura 1.5 muestra los artículos aceptados y rechazados por cada fuente de datos, después de eliminar los duplicados. Por tanto, una vez finalizado el proceso de selección de trabajos, se aceptaron 108 artículos de Scopus y 32 de IEEE Xplore.

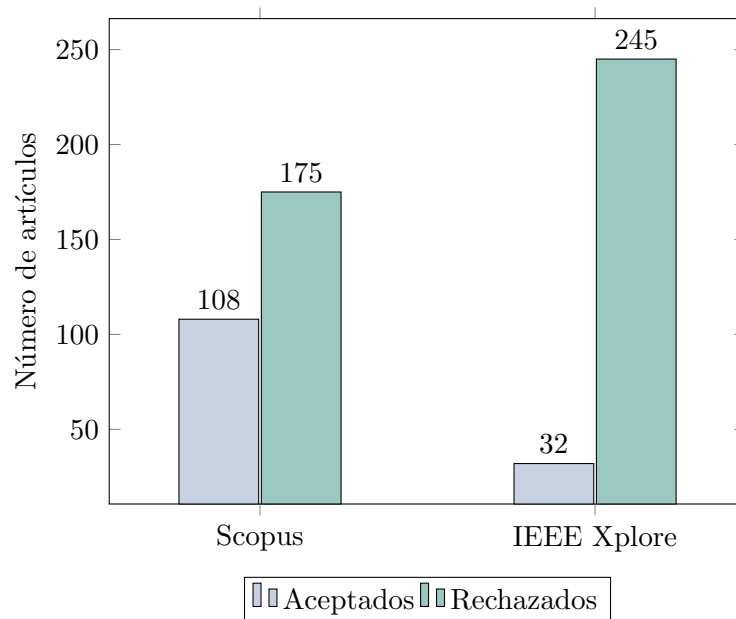


Figura 1.5: Número de artículos aceptados y rechazados por fuente

Además, la Figura 1.6 muestra el número total de artículos por año después de aplicar los criterios de inclusión y exclusión. Tal y como se puede observar, se ha producido un incremento significativo de estudios relacionados en los últimos años.

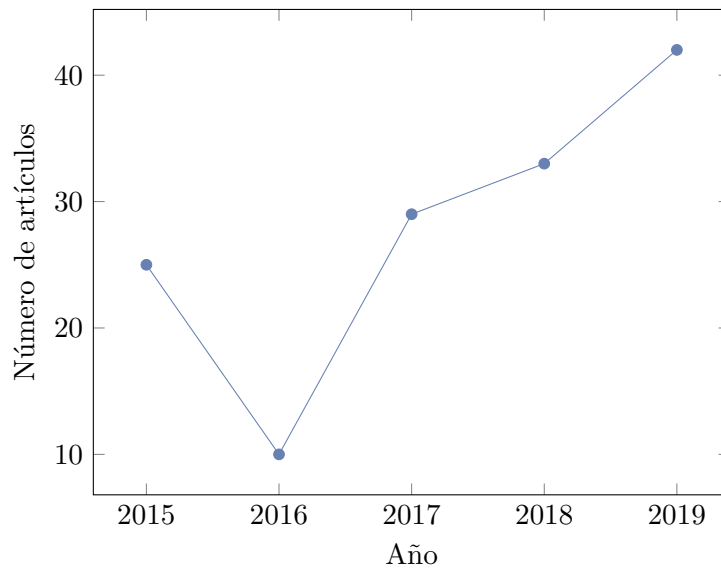


Figura 1.6: Número total de artículos por año

1.4. Estructura de la Tesis

Una vez realizada y analizada la revisión sistemática de la literatura se tiene una visión general del tema de investigación que se está tratando. Teniendo en cuenta el objetivo principal y los objetivos específicos definidos en la sección 1.2, presentaremos la organización de los estudios realizados, en base a una serie de preguntas de investigación (RQ), tal y como se muestra en la Figura 1.7.

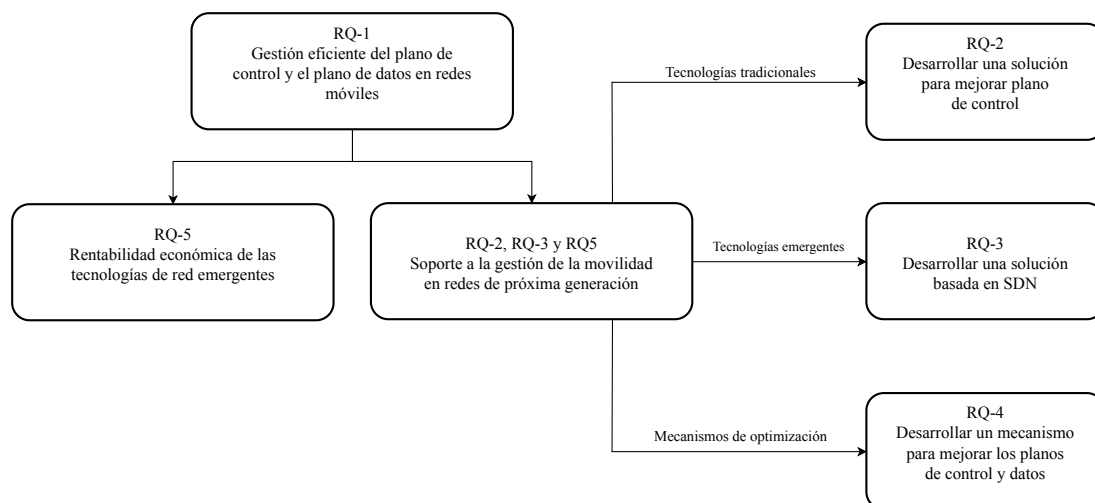


Figura 1.7: Organización de la Tesis en base a las preguntas de investigación

Estas RQ complementarán a las establecidas en la SLR, pero estarán centradas en nuestros estudios y serán definidas específicamente para conseguir el objetivo principal de la investigación. Por tanto, se han definido cinco preguntas de investigación que se irán respondiendo con el desarrollo de la Tesis:

- **RQ-1:** ¿Cómo se pueden modificar las redes móviles de próxima generación para llevar a cabo una gestión eficiente del plano de control y del plano de datos?
- **RQ-2:** Haciendo uso de mecanismos tradicionales, ¿cómo se puede mejorar el plano de control para los protocolos de gestión de la movilidad?
- **RQ-3:** ¿Cómo se pueden aplicar tecnologías emergentes como SDN para mejorar la gestión de la movilidad en redes móviles de próxima generación?
- **RQ-4:** ¿Cómo se pueden aplicar mecanismos de optimización para mejorar los planos de control y de datos cuando se ofrece soporte a la gestión de la movilidad en redes móviles de próxima generación?
- **RQ-5:** ¿Cuánta rentabilidad económica ofrecen las tecnologías de red emergentes?

Tal y como se muestra en la Figura 1.8, cada pregunta de investigación se corresponde con un capítulo determinado. Además, en la Figura 1.8 también se muestra la relación entre los distintos capítulos del presente documento.

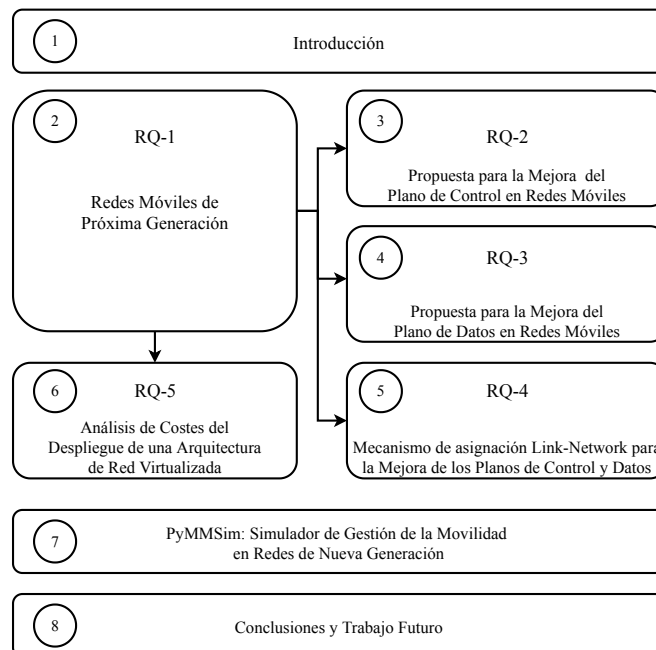


Figura 1.8: Organización de la Tesis y relación entre los capítulos

En definitiva, la investigación llevada a cabo se estructura tal y como se describe a continuación. El Capítulo 1 presenta una contextualización de la investigación desarrollada, planteando la problemática principal que afecta a las redes móviles de próxima generación, los objetivos que se pretenden alcanzar, los antecedentes y las aportaciones realizadas.

Los trabajos relacionados se introducen en el Capítulo 2, proporcionando una visión general de la redes móviles de próxima generación. Se describen los principales enfoques de gestión de la movilidad y su evolución, introduciendo la necesidad de nuevos mecanismos que solucionen los problemas y limitaciones que presentan estos enfoques en la actualidad.

Los siguientes tres capítulos contienen las aportaciones principales de esta Tesis.

En el Capítulo 3, se implementa un nuevo mecanismo de gestión de la movilidad distribuida que presenta beneficios desde el punto de vista del plano de control, consiguiendo reducciones significativas del tráfico de señalización.

El Capítulo 4 describe el segundo mecanismo propuesto en esta Tesis. En este caso, se ha desarrollado una solución de gestión de la movilidad distribuida que hace uso de los beneficios de tecnologías emergentes como SDN, con el principal objetivo de proporcionar mejoras en el plano de datos.

A continuación, en el Capítulo 5 se ha desarrollado un mecanismo que busca la optimización en la asignación de las estaciones bases con la red de acceso, consiguiendo importantes mejoras en el plano de control y el plano de datos en redes móviles de próxima generación.

El Capítulo 6 muestra un análisis y evaluación de costes del despliegue de tecnologías de red emergentes, estudiando la viabilidad económica de las mismas.

El Capítulo 7 presenta el diseño y la implementación del simulador de gestión de la movilidad desarrollado para obtener los resultados de simulación presentados en esta Tesis.

Por último, el Capítulo 8 resume la investigación realizada y resultados obtenidos a modo de conclusiones, así como algunas propuestas de trabajo futuro.

1.5. Contribuciones

En esta sección se presentan las principales aportaciones realizadas en esta Tesis.

- Un estudio de oportunidades y retos en las redes móviles de próxima generación, prestando especial atención al proceso de soporte a la movilidad. Por ello, se presentan los diferentes protocolos de gestión de la movilidad, así como su evolución, desde los enfoques de red centralizados a los enfoques distribuidos, analizando los problemas y limitaciones de cada uno de ellos.
- Un mecanismo de gestión de la movilidad, llamado *Tunneling Extension to Distributed Mobility Management* (TE-DMM), ha sido propuesto para mejorar el rendimiento en arquitecturas distribuidas. Esta solución permite llevar a cabo una gestión eficiente del plano de control, reduciendo en gran medida el tráfico de señalización. Se proporciona un análisis y una evaluación de manera analítica y experimental, obteniendo resultados de costes de señalización, pérdida de paquetes, costes de procesamiento y latencia de *handover*. Además, se ha implementado el mecanismo sobre el *kernel* de Linux para poder evaluarlo convenientemente en un entorno real.
- Una propuesta de gestión de la movilidad basada en *software* que combina SDN con DMM, proporcionando flexibilidad, escalabilidad y fiabilidad a las redes móviles de próxima generación y mejorando el rendimiento del plano de datos. Para ello, nuestra propuesta hace uso de las capacidades de SDN con el objetivo de gestionar la movilidad como un servicio. Esta propuesta ha sido evaluada en términos de coste de señalización, coste de entrega de paquetes y latencia de *handover*. Además, se ha llevado a cabo la implementación del mecanismo sobre un entorno de red real para validarlo experimentalmente.
- Una estrategia de asociación entre estaciones base y la red de acceso que proporciona beneficios al rendimiento de los protocolos de gestión de la movilidad, desde dos perspectivas: plano de control y plano de datos. Este mecanismo es evaluado y comparado con respecto a otros en términos de costes de movilidad: costes de señalización y costes de entrega de paquetes.
- Desarrollo de un simulador en Python para evaluar los distintos mecanismos propuestos, bajo diferentes condiciones de movilidad, diferentes modelos de tráfico y diferentes topologías de red.
- Una evaluación de costes relacionados con el despliegue de tecnologías de red emergentes como SDN o NFV, analizando la viabilidad económica. Además, se propone un modelo de costes basado en costes de capital (CAPEX) y costes de operación (OPEX) que permita obtener una evaluación comparativa fiable entre arquitecturas de red tradicionales y arquitecturas de red virtualizadas.

1.6. Publicaciones

A continuación se muestra una relación de las publicaciones realizadas durante el desarrollo de esta investigación.

Revistas

- Jesús Calle-Cancho, Javier Carmona-Murillo, José-Luis González-Sánchez and David Cortés-Polo. A Novel Link-Network Assignment Strategy to Enhance the Performance of Mobility Management Protocols. 2020. Under review.
- Jesús Calle-Cancho, David Cortés-Polo, José-Luis González-Sánchez, Luis-Ignacio Jiménez-Gil and Javier Carmona-Murillo. Analytical and Experimental Evaluation of a Novel Mechanism to Improve the Control Plane in Next-Generation Mobile Networks. *Electronics* 2020, 9, 417. JCR: 2.412. URL: <https://doi.org/10.3390/electronics9030417>
- Jesús Calle-Cancho, José-Manuel Mendoza-Rubio, José-Luis González-Sánchez, David Cortés-Polo and Javier Carmona-Murillo. Towards a software-based mobility management for 5G: An experimental approach for flattened network architectures. *Computer Science and Information Systems* 2020, 17-1, pp. 51-70. JCR: 0.927. URL: <https://doi.org/10.2298/CSIS181101016C>
- David Cortés-Polo, Luis-Ignacio Jiménez-Gil José-Luis González-Sánchez and Jesús Calle-Cancho. A Methodology for Network Analysis to Improve the Cyber-Physicals Communications in Next-Generation Networks. *Sensors* 2020, 20, 2247. JCR: 3.275. URL: <https://doi.org/10.3390/s20082247>
- David Cortés-Polo, Luis-Ignacio Jiménez-Gil, Jesús Calle-Cancho and José-Luis González-Sánchez. A Novel Methodology Based on Orthogonal Projections for a Mobile Network Data Set Analysis. *IEEE Access* 2019, 7, pp. 158007-158015. JCR: 3.745. URL: <https://doi.org/10.1109/ACCESS.2019.2949804>
- Joaquín Espinosa-García, Jesús Calle-Cancho and José-Carlos Corchado-Romo. QCT Study of the Vibrational and Translational Role in the H + C₂H₆ Reactions. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling* 2019. 138-10:116. JCR: 1.498. URL: <https://doi.org/10.1007/s00214-019-2504-4>
- José-Manuel Mendoza-Rubio, Javier Carmona-Murillo, José-Luis González-Sánchez, Jesús Calle-Cancho and David Cortés-Polo. SR-DMM: An SDN-Based DMM Solution For Future Mobile Networks. *IEEE*

Latin America Transactions 2019, 17-5, pp. 734-741. JCR: 0.782. URL: <https://doi.org/10.1109/TLA.2019.8891941>

- Jesús Calle-Cancho, Javier Carmona-Murillo, David Cortés-Polo, José-Luis González-Sánchez and Francisco-Javier Rodríguez-Pérez. TE-DMM: A Proposal to Improve the Control Plane in PMIPv6-based DMM Networks. *IEEE Latin America Transactions* 2015, 13-9, pp. 3149-3155. JCR: 0.436. URL: <https://doi.org/10.1109/TLA.2015.7350071>
- David Cortés-Polo, Jesús Calle-Cancho, Javier Carmona-Murillo and José-Luis González-Sánchez. Future Trends in Mobile-Fixed Integration for Next Generation Networks: Classification and Analysis. *International Journal of Vehicular Telematics and Infotainment Systems* 2017, 1. URL: <https://www.doi.org/10.4018/IJVTIS.2017010103>
- Javier Carmona-Murillo, David Cortés-Polo, Jesús Calle-Cancho, José-Luis González-Sánchez, Francisco-Javier Rodríguez-Pérez. Analytical and Experimental Evaluation of Handovers in IPv6 Mobility Management Protocols. *Network Protocols and Algorithms* 2016, 8-1. URL: <https://doi.org/10.5296/npa.v8i1.8725>

Congresos

- Jesús Calle-Cancho, Javier Carmona-Murillo, David Cortés-Polo y José-Luis González-Sánchez. Análisis Tecno-Económico del Despliegue de una Arquitectura de Red Virtualizada. *IADIS International Conference WWW/Internet* 2019, pp. 269-274. ISBN 978-989-8533-96-8.
- Jesús Calle-Cancho, David Cortés-Polo, Javier Carmona-Murillo, José-Luis González-Sánchez y Francisco-Javier Rodríguez-Pérez. Análisis de costes del despliegue de una arquitectura de red basada en SDN/NFV. *XIV Jornadas de Ingeniería Telemática* 2019, pp. 215-221. ISBN: 978-84-09-21112-8.
- David Cortés-Polo, Felipe Lemus-Prieto, Jesús Calle-Cancho, Luis-Ignacio Jiménez-Gil and José-Luis González-Sánchez. Gestión de seguridad de las comunicaciones para entornos de HPC en centros de supercomputación. *Jornadas Sarteco* 2019, pp. 437-443. ISBN: 978-84-09-12127-4.
- Luis-Ignacio Jiménez-Gil, Jesús Calle-Cancho, David Cortés-Polo and José-Luis González-Sánchez. A New Role for Supercomputing Centers in Open Science. *IBERGRID* 2018.

- Javier Carmona-Murillo, David Cortés-Polo, Jesús Calle-Cancho, José-Luis González-Sánchez, Francisco-Javier Rodríguez-Pérez. Experimental evaluation of mobility management protocols for multimedia communications. *XII Jornadas de Ingeniería Telemática* 2015, pp. 293-300. ISBN: 978-84-606-8609-5.
- Jesús Calle-Cancho, David Cortés-Polo, Javier Carmona-Murillo, José-Luis González-Sánchez y Francisco-Javier Rodríguez-Pérez. Análisis y evaluación experimental de protocolos de gestión de la movilidad basados en IPv6. *IADIS International Conference WWW/Internet* 2014. pp. 35-42. ISBN: 978-989-8704-12-2.
- David Cortés-Polo, Javier Carmona-Murillo, Jesús Calle-Cancho, José-Luis González-Sánchez, Francisco-Javier Rodríguez-Pérez. Evaluación experimental de la gestión de la movilidad en redes IPv6. *Jornadas Técnicas de RedIRIS* 2014.

Capítulos de libro

- José-Luis González-Sánchez, Jesús Calle-Cancho, David Cortés-Polo, Luis-Ignacio Jiménez-Gil and Alfonso López-Rourich. High Performance Computing, Big Data and Cloud Computing: The Perfect De Facto Trio or Converging Technological Mantras. pp. 165-182, IGI Global. 2019. ISBN: 9781522570745. URL: <https://www.doi.org/10.4018/978-1-5225-7074-5.ch009>

1.7. Participación en proyectos de investigación

A continuación se muestra una relación de los proyectos de investigación en los que se ha participado durante el desarrollo de la Tesis.

- SmartNet5G: Desarrollo de nuevos mecanismos de gestión en redes programables de próxima generación (IB18003).
 - Investigador principal: Javier Carmona Murillo
 - Entidad financiadora: Consejería de Economía e Infraestructuras
 - Nº de investigadores: 6
 - Duración: 09/02/2019 - 17/07/2022
 - Cuantía: 139 991.50 €

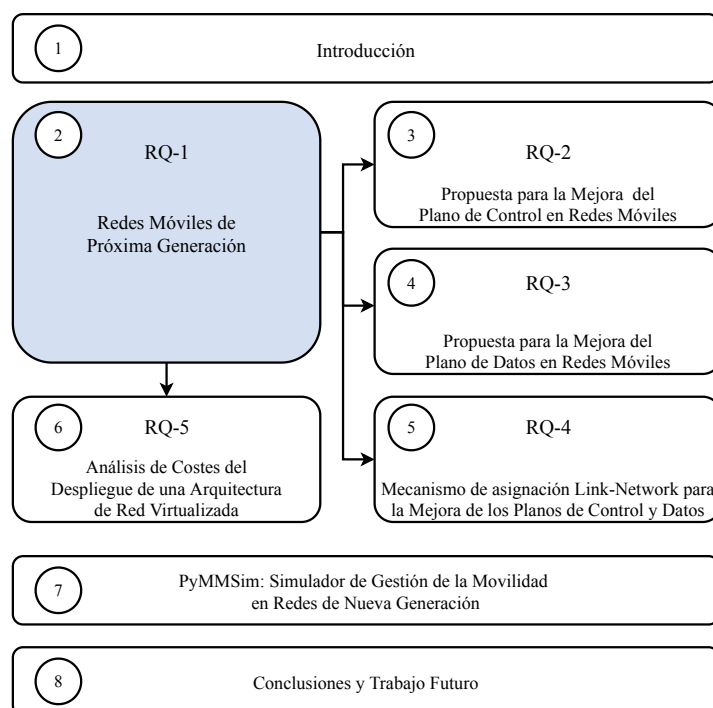
- CénitS-CPD: Adquisición de equipamiento científico tecnológico.
 - Investigador principal: José Luis González Sánchez
 - Entidad financiadora: Consejería de Economía, Ciencia y Agenda Digital
 - Nº de investigadores: 7
 - Duración: 25/04/2020 - 31/12/2020
 - Cuantía: 1 000 000.00 €
- Optimización de algoritmos y aplicaciones paralelas en sistemas heterogéneos mediante el uso combinado de modelos formales de cómputo y comunicaciones (IB16118).
 - Investigador principal: Juan Antonio Rico Gallego
 - Entidad financiadora: Consejería Economía, Competitividad e Innovación
 - Nº de investigadores: 7
 - Duración: 03/06/2017 - 02/06/2020
 - Cuantía: 85 800.00 €
- HPCénitS: Adquisición de equipamiento científico tecnológico.
 - Investigador principal: José Luis González Sánchez
 - Entidad financiadora: Consejería de Economía e Infraestructuras
 - Nº de investigadores: 4
 - Duración: 08/05/2019 - 31/12/2019
 - Cuantía: 712 319.00 €
- Desarrollo de una red de infraestructuras federadas para la generación de servicios de virtualización de puestos de trabajo (SOE/P3/E804).
 - Investigador principal: Carlos Redondo Gil
 - Entidad financiadora: FEDER
 - Nº de investigadores: 35
 - Duración: 02/10/2014 - 30/06/2015
 - Cuantía: 859 403.00 €
- Modernización e Innovación Tecnológica con base TIC en Sectores Estratégicos y Tradicionales (0606_MITTIC_4.E).
 - Investigador principal: Sergio Velázquez Vioque

- Entidad financiadora: Fondo Europeo de Desarrollo Regional
- Nº de investigadores: 30
- Duración: 01/05/2013 - 30/09/2015
- Cuantía: 1 521 964.84 €

Capítulo 2

Redes Móviles de Próxima Generación

En este capítulo se proporciona una visión general de las redes móviles de próxima generación y se presentan los diferentes protocolos de gestión de la movilidad, así como su evolución, desde los enfoques centralizados a los enfoques distribuidos. Además, se exponen los trabajos y tecnologías relacionadas con la Tesis, sentando las bases de la misma e introduciendo la necesidad de nuevos mecanismos que solucionen las limitaciones que afectan a los actuales protocolos de gestión de la movilidad.



El capítulo se organiza tal y como se presenta a continuación:

- La sección 2.1 introduce la necesidad de mecanismos que gestionen la movilidad de manera eficiente en redes móviles de próxima generación.
- La sección 2.2 muestra la evolución de las arquitecturas de red, prestando especial atención a la separación de los planos de control y de datos.
- La sección 2.3 describe los principales protocolos de gestión de la movilidad centralizada estandarizados por el IETF, analizando sus principales limitaciones. También se presentan los mecanismos de gestión de la movilidad distribuida más representativos que intentan solucionar los problemas que afectan a los enfoques centralizados.

Además se muestra una comparativa entre las soluciones de gestión de la movilidad, estableciendo las principales diferencias entre ellas. Finalmente, se presenta un análisis cualitativo de los trabajos de investigación relacionados con la gestión de la movilidad que han sido publicados en los últimos años, para identificar las métricas que nos permitan evaluar el rendimiento de las contribuciones presentadas en esta tesis.

2.1. Introducción

En los últimos años, el tráfico de datos móviles ha experimentado un crecimiento exponencial debido al continuo desarrollo de las redes de comunicaciones móviles. Estos avances, unidos a la aparición de multitud de servicios ofertados por los operadores de red, han impulsado la creación de nuevos mecanismos y arquitecturas de gestión de redes eficientemente, capaces de soportar tecnologías emergentes y nuevos servicios y aplicaciones móviles.

Además, estudios recientes estiman que se producirá un crecimiento importante de dispositivos interconectados entre 2017 y 2022, incrementándose en un 37%. Se prevé que el tráfico global de datos móviles se incremente siete veces durante el periodo 2017-2022, estimándose un consumo mensual de 77.5 exabytes en 2022 [18, 19]. Además, las previsiones indican que en 2023 habrá más de mil millones de clientes de las futuras redes 5G.

Esta nueva generación de comunicaciones móviles (5G) no solo se plantea como una evolución de las redes 3G y 4G, sino que soporta una gran variedad de servicios y aporta mejoras desde el punto de vista de la banda ancha móvil, proporcionando

alta fiabilidad y baja latencia a las comunicaciones. La arquitectura funcional de las anteriores generaciones de tecnologías de comunicaciones era de naturaleza centralizada, ya que su principal objetivo se basaba en proporcionar soporte a las comunicaciones *humano-humano* con un limitado número de servicios (banda de ancha móvil, servicio de voz y SMS) [20]. Sin embargo, las redes móviles están evolucionando y 5G se plantea como una tecnología que dará respuesta a retos complejos, exigiendo un avance tecnológico significativo con respecto al estado actual de la tecnología. Por lo tanto, se espera que las redes de próxima generación sean capaces de proporcionar nuevos servicios de acuerdo a las demandas específicas de los usuarios, ya que las nuevas formas de comunicación son muy diferentes de las comunicaciones tradicionales en términos de fiabilidad, latencia, flexibilidad y eficiencia. Estos estrictos requisitos de rendimiento deben ser considerados en el diseño de arquitecturas de redes móviles de próxima generación, pero los principios de las arquitecturas 3G y 4G no son suficientes para cubrir los requisitos de las redes 5G [21].

Para abordar servicios y aplicaciones diversificados, el 3GPP ha clasificado estos servicios y aplicaciones 5G en tres categorías principales: *Ultra-Reliable and Low-latency Communication* (URLLC), *Massive Internet of Things* (MIoT) y *enhanced Mobile Broadband* (eMBB). Estas categorías admiten servicios y aplicaciones con estrictos requisitos de rendimiento, como vehículos sin conductor y vehículos aéreos no tripulados, atención médica remota y cirugía robótica, ciudades inteligentes e industria, etc [22].

Por lo tanto, URLLC se convierte en un reto importante para los entornos de red 5G, ya que es muy difícil producir mejoras de alta fiabilidad y baja latencia simultáneamente [23]. Con el objetivo de conseguir estos estrictos requisitos de rendimiento, una práctica habitual es incrementar los recursos de transmisión, la redundancia de señalización e, incluso, desplegar más estaciones base en una determinada región, dando lugar a redes ultra-densas [22]. Sin embargo, esta densificación de celdas a veces incrementa la latencia y puede producir degradación de la calidad de servicio (QoS, *Quality of Service*) y la calidad de experiencia (QoE, *Quality of Experience*) [24].

En este contexto en el que multitud de servicios, con diferentes requerimientos, comparten una red heterogénea, resulta imprescindible desarrollar mecanismos que permitan asignar, de forma óptima, los recursos de la red física a las necesidades de conectividad de los usuarios. Por tanto, se han impulsado una serie de tecnologías eficientes para la gestión de red que proporcionan numerosos beneficios al ecosistema 5G, sentando las bases del mismo [25]. Entre estas tecnologías emergentes destacan SDN, NFV y MEC, con el objetivo general de alcanzar el rendimiento necesario, la escalabilidad y la flexibilidad esperada.

SDN [26, 27] es un paradigma de red que ha atraído la atención de la academia y la industria en los últimos años y, surgió como una arquitectura para gestionar las redes tradicionales; las cuales presentan limitaciones importantes, son complejas y difíciles de administrar [28, 29]. Por lo tanto, SDN ha supuesto un cambio significativo en la manera en la que se diseñan y administran las redes, superando las limitaciones que afectan a las infraestructuras de red actuales.

Las arquitecturas SDN están basadas en la separación del plano de datos del plano de control [30]. La inteligencia y estado de la red están centralizados lógicamente, y la infraestructura de red subyacente se abstrae de las aplicaciones [31]. Por tanto, la integración vertical desaparece al separar la lógica de control (plano de control) de los dispositivos físicos de red que envían el tráfico (plano de datos). A diferencia de las arquitecturas de red tradicionales, el paradigma SDN plantea que el control de la red pase a estar en un controlador centralizado, que se encargará de tomar decisiones sobre el plano de datos de cada uno de los elementos de la red. La Figura 2.1 muestra una comparativa entre las redes tradicionales, donde la lógica está distribuida entre los dispositivos de la red, y el paradigma SDN donde la lógica está centralizada.

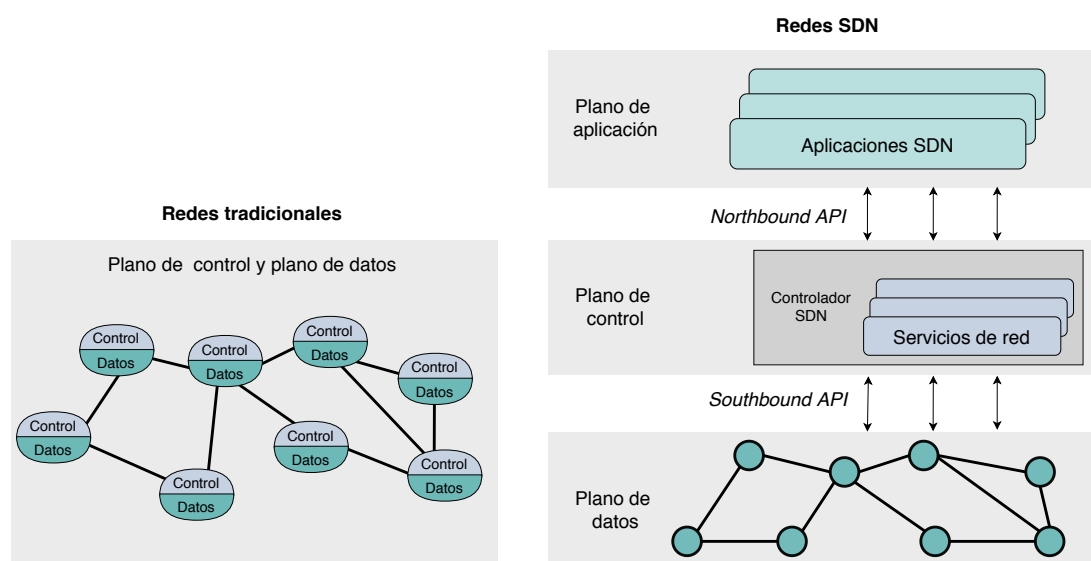


Figura 2.1: Comparativa entre una arquitectura de red tradicional y una arquitectura de red SDN

De esta forma, la inteligencia de la red se mantiene en un dispositivo que toma las decisiones oportunas sobre el estado completo de la red [32]. A través de la interfaz superior (*Northbound API*), el controlador proporciona una visión global de la red a las aplicaciones, mientras que a través de la interfaz inferior (*Southbound API*) se comunica con los dispositivos de red utilizando alguno de los protocolos estandarizados para esta tarea. Por lo tanto, el control de la red puede ser programado directamente a través

de una interfaz abierta. Entre las más utilizadas se encuentran OpenFlow [33], ForCES [34], etc. De esta forma, la infraestructura es considerada como una red de reenvío de paquetes (plano de datos) que se puede programar por los administradores de red para ejecutar aplicaciones SDN que sean capaces de optimizar servicios específicos.

Tal y como muestra la Figura 2.1, la arquitectura funcional de SDN se divide en tres planos principalmente. La capa de infraestructura está formada por el equipamiento físico de red (*switches*) y conforma el plano de datos. Por otro lado, la capa de control engloba a los controladores que facilitan la configuración de flujos y el encaminamiento en la red. Finalmente, en la capa de aplicación se contemplan las aplicaciones funcionales, que pueden ser consideradas como servicios específicos basados en SDN.

Además, la aparición del paradigma de NFV ha sido muy importante desde el punto de vista del aprovisionamiento de servicios de telecomunicaciones. Este paradigma tiene como principal objetivo la separación de las funciones de red de los dispositivos físicos en los cuales se ejecutan [35]. Con NFV, un servicio se puede descomponer en un conjunto de funciones de red virtuales (VNF), que son implementadas mediante *software*. Una VNF puede ser instanciada en diferentes ubicaciones de la red sin requerir la instalación de nuevos dispositivos físicos, proporcionando funcionalidades de red específicas, tales como balanceo de carga, cortafuegos, redes privadas virtuales (VPN), etc. NFV ofrece la posibilidad de que estas instancias de funciones virtuales sean compartidas entre distintos clientes sobre una misma infraestructura de red [36], permitiendo una implementación automatizada y escalable de las mismas.

Tal y como se puede observar en la Figura 2.2, la arquitectura de NFV se compone de tres bloques principales que se describen a continuación [37]:

- Infraestructura física: es la base general de la arquitectura NFV que contiene el *hardware* necesario para soportar los recursos virtuales de la capa superior.
- Infraestructura virtual: contiene el *software* que hace posible la virtualización y los propios recursos virtuales. Este bloque interactúa tanto con la infraestructura física como con la capa de funciones virtuales de red, incluyendo las funcionalidades del despliegue, gestión y ejecución de VNF.
- Funciones virtuales de red: está formada por las funciones de red virtuales que son desplegadas sobre la infraestructura virtual.

Los operadores de redes móviles tienen la necesidad de reducir los costes de capital (CAPEX) y los costes de operación (OPEX) de sus infraestructuras. Por ello, NFV

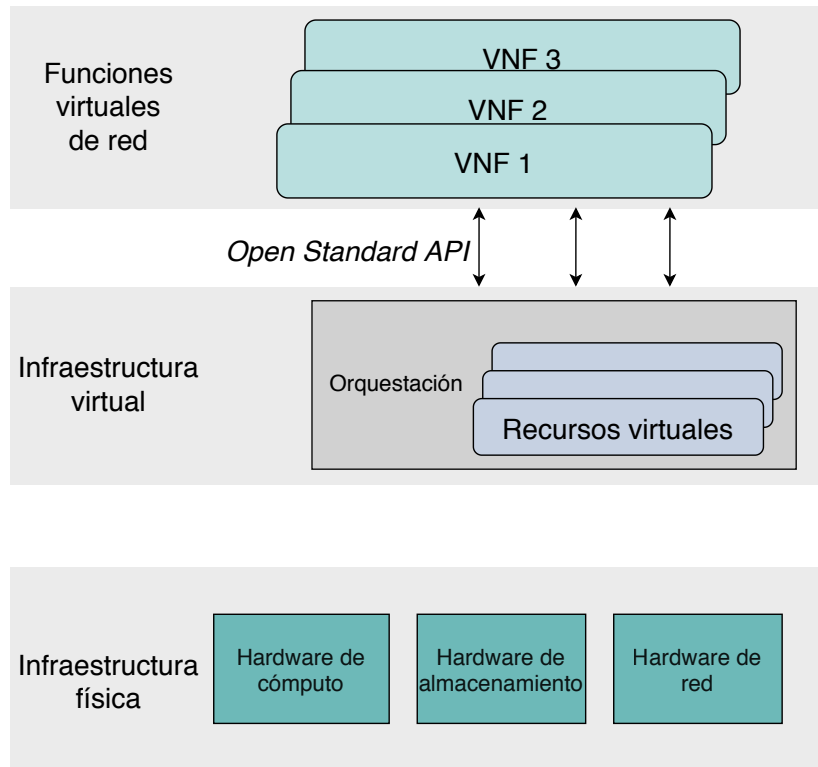


Figura 2.2: Arquitectura de red de NFV

es considerada como una tecnología que tiene el potencial de proporcionar reducciones significativas de CAPEX y OPEX, además de facilitar y flexibilizar el despliegue de nuevos servicios con más agilidad [38]; permitiendo alcanzar los requisitos de baja latencia y alta fiabilidad requerida por los servicios que se ofrecerán en las futuras redes 5G [39]. En este sentido, desde ETSI se siguieron haciendo esfuerzos para mejorar las comunicaciones con estrictos requisitos de rendimiento y, por ello, desde uno de sus grupos de trabajo (ISG, *Industry Specification Group*), se planteó el concepto de MEC [40], con el objetivo de llevar a cabo una integración eficiente y transparente de las funcionalidades del paradigma *Cloud Computing* hacia la red móvil [41], dando lugar a una plataforma donde las capacidades de computación y almacenamiento se distribuyen por el borde de la red. De esta manera, se reduce la distancia de comunicación y, por consiguiente, los posibles retardos entre la red móvil y los usuarios finales [23]. Este planteamiento puede beneficiarse de la propia infraestructura de virtualización NFV, buscando la integración óptima de la red móvil y las funciones virtuales de red [24].

Por otro lado, 5G PPP (*5G Infrastructure Public Private Partnership*), a través de uno de sus grupos de trabajo (*Architecture Working Group*¹) ha sentado las bases

¹<https://5g-ppp.eu/5g-ppp-work-groups/>

de los casos de uso emergentes en las redes 5G [42], donde uno de los aspectos clave es el soporte a la movilidad, abordando los requisitos de dichos casos de uso en diferentes ámbitos, como automóviles autónomos y vehículos aéreos no tripulados, atención médica remota, robótica industrial, ciudades inteligentes, etc. Por lo tanto, las redes móviles 5G imponen estrictos desafíos de gestión de la movilidad, ya que la mayoría de los servicios, aplicaciones y casos de uso contemplan rigurosas restricciones de rendimiento [43]. En este entorno, se necesitan mecanismos efectivos de gestión de la movilidad que proporcionen soporte a usuarios móviles con distintos perfiles de movilidad [44], desde 0 km/h (sensores estáticos) hasta 300 km/h (usuarios en trenes de alta velocidad).

Estos mecanismos proporcionan soporte a la movilidad de manera transparente en entornos de red heterogéneos, sin interrupciones a nivel de red, manteniendo las comunicaciones activas durante el movimiento de los usuarios móviles entre distintas redes de acceso. Estos protocolos se pueden clasificar en enfoques centralizados y distribuidos. Los protocolos de gestión de movilidad centralizada (CMM) [45] introducen un agente, denominado ancla de movilidad, que es el encargado de gestionar y mantener la comunicación activa durante el movimiento de los usuarios. Sin embargo, estas arquitecturas CMM tienen ciertas limitaciones y problemas de fiabilidad, escalabilidad y encaminamiento sub-óptimo, entre otros. Por ello, se han propuesto nuevos enfoques de gestión de movilidad distribuida (DMM) [8, 46], en los que las anclas de movilidad se ubican más cerca del usuario, con el objetivo de obtener una red más plana (*flatter network*), reduciendo los cuellos de botella que afectan a las redes móviles actuales debido al crecimiento del tráfico de datos.

Sin embargo, estos mecanismos de gestión de movilidad introducen sobrecarga de señalización. Este aspecto, junto a la densificación celular producida por 5G, hará que aumente el tráfico total de señalización, degradando los requisitos de QoS y QoE.

Por tanto, los operadores buscan soluciones innovadoras para la optimización de los procedimientos de gestión de la movilidad dentro de la arquitectura de red 5G. Los operadores de red móvil tendrán que tener en cuenta los aspectos mencionados anteriormente al planificar las arquitecturas de próxima generación, con el objetivo de mejorar el rendimiento de la red y diseñar infraestructuras capaces de soportar los requisitos de los casos de uso emergentes en 5G. Además, estas asignaciones estratégicas entre las estaciones base y los nodos de acceso mejorarán significativamente el rendimiento de los protocolos de movilidad.

2.2. Evolución de las arquitecturas de red

Las arquitecturas de red han evolucionado en los últimos años hacia una separación completa del plano de datos y el plano de control. Actualmente, esta separación entre planos también ha sido considerada como un factor clave en el diseño de redes de próxima generación [47]. Con este concepto, las funciones del plano de control se pueden implementar por *software* en una plataforma *cloud* para conseguir la flexibilidad y escalabilidad necesarias en estos entornos a medida que aumenta el tráfico de señalización. Además, las funciones del plano de datos se pueden desplegar en dispositivos de red *hardware* simplificados y de alta velocidad, optimizados para el reenvío de paquetes.

En la actualidad, parece que la separación de los planos de datos y control ha surgido ligada a la tecnología SDN, pero esto no es del todo cierto [48]. SDN es parte de un proceso evolutivo de esfuerzos continuados para conseguir una separación completa y eficiente entre planos. Este concepto ha sido considerado desde mediados de los años 90 y, por tanto, no se limita a SDN.

La Figura 2.3 muestra una línea temporal con la evolución y los hitos más importantes conseguidos en relación a la separación entre el plano de control y el plano de datos, que se explicarán con más detalle a continuación.

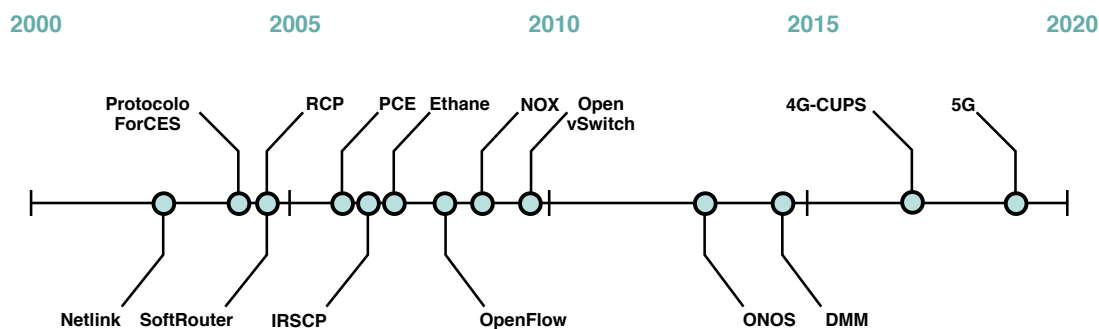


Figura 2.3: Evolución de la separación entre el plano de datos y el plano de control

2.2.1. Primeros esfuerzos de separación de los planos de control y de datos

Desde la primera década del siglo XXI, se ha producido un gran incremento del volumen de tráfico de datos, llevando a los operadores de red a buscar mecanismos de gestión más eficientes para mejorar el rendimiento global de la red como, por ejemplo, la ingeniería de tráfico. Inicialmente, la comunidad investigadora junto a los operadores

exploraron enfoques pragmáticos a corto plazo, que estuvieran basados en estándares o que se pudieran implementar de manera inminente utilizando los protocolos existentes.

Específicamente, los *routers* y *switches* convencionales presentaban una clara integración entre el plano de control y el plano de datos. Este acoplamiento dificultaba las tareas de administración, gestión y operación de la red. Por lo tanto, para superar estas limitaciones, se comenzaron a buscar alternativas persiguiendo la separación de las funciones del plano de control de las del plano de datos. Estos esfuerzos dieron lugar a dos innovaciones principalmente, que se muestran a continuación:

- *Interfaces* abiertas entre el plano de control y el plano de datos, como la interfaz ForCES (*Forwarding and Control Element Separation*) [49] estandarizada por el IETF y la interfaz Netlink para la funcionalidad de reenvío de paquetes a nivel de *kernel* en Linux [50].
- Control lógico de la red centralizado: arquitecturas *Routing Control Platform* (RCP) [51, 52] y *SoftRouter* [53], además del protocolo *Path Computation Element* (PCE) [54] estandarizado por el IETF.

Estas innovaciones fueron impulsadas por la creciente demanda de la industria tecnológica para administrar el encaminamiento en los proveedores de servicios de red. Además, algunas de las primeras propuestas para la separación de los planos de datos y control se llevaron a cabo en entornos académicos relacionados con redes activas [55] y redes ATM (*Asynchronous Transfer Mode*) [56, 57]. Sin embargo, estas propuestas eran muy diferentes a los proyectos mencionados anteriormente, los cuales se centraron en problemas importantes en la gestión de la red, prestando especial atención a la innovación por y para los administradores de red (en lugar de los usuarios finales e investigadores), programabilidad en el plano de control y visibilidad y control en toda la red (en lugar de la configuración a nivel de dispositivo).

Posteriormente, las aplicaciones de administración de red incluyeron la selección de mejores rutas de red en función de la carga de tráfico actual, minimizando las interrupciones transitorias y proporcionando más control sobre el flujo de tráfico. Numerosas aplicaciones de control se ejecutaron sobre redes en operación utilizando los mismos dispositivos de red físicos, entre las que destaca *Intelligent Route Service Control Point* (IRSCP) [58], implementado para ofrecer servicios de valor añadido para clientes de redes privadas virtuales en la red troncal de nivel 1 de AT&T (*American Telephone and Telegraph*).

En años sucesivos, para ampliar la visión de la separación del plano de control y el plano de datos, los investigadores comenzaron a explorar nuevas arquitecturas que

podieran llevar a cabo un control centralizado de la red. Varios grupos de investigación procedieron a diseñar y construir sistemas que aplicaron este enfoque de alto nivel a nuevas áreas de aplicación [59, 60], más allá del control de rutas. En particular, en el proyecto Ethane [59] se creó una solución de nivel de flujo lógicamente centralizada para el control de acceso en redes empresariales. Ethane reduce los *switches* a tablas de flujo que son gestionadas por un controlador con unas determinadas políticas. Este proyecto y su despliegue operativo en el departamento de informática de Stanford se convirtieron en la base de la API original de OpenFlow.

En definitiva, los intentos iniciales de separar los planos de control y de datos fueron relativamente pragmáticos, pero representaron una desviación conceptual significativa del acoplamiento que se producía en las arquitecturas de red tradicionales. Los esfuerzos llevados a cabo para separar el control de la red y el plano de datos dieron como resultado varios conceptos que han sido utilizados en diseños posteriores, tales como SDN.

2.2.2. Redes programables: OpenFlow

A partir del año 2008, tal y como muestra la Figura 2.3, los trabajos relacionados con la programabilidad de red y la separación de los planos de control y de datos cobraron mayor importancia, siendo impulsados por la aparición del protocolo OpenFlow, que logró un equilibrio entre la visión de las redes completamente programables y el pragmatismo que permitiría el despliegue en el mundo real. OpenFlow habilitó más funciones de gestión en los controladores de red y fue implementado sobre los dispositivos de red existentes en aquel momento. Aunque confiar en el *hardware* existente generaba una serie de limitaciones de flexibilidad, OpenFlow se podía implementar casi de forma directa, permitiendo un movimiento pragmático hacia SDN y la consiguiente separación del plano de control y el plano de datos.

La creación de la API OpenFlow [61] fue seguida al poco tiempo por el diseño de plataformas de controladores como NOX [62] que permitieron la creación de nuevas aplicaciones de control. Seguidamente, surgió Open vSwitch [63], una implementación de código abierto, diseñado para ser utilizado como un *switch* virtual o *software switch*, que encapsula el tráfico destinado a máquinas virtuales que se ejecutan en otros servidores diferentes [64]. Un controlador centralizado instala las reglas en estos *switches* virtuales para controlar la encapsulación de los paquetes y actualizar las reglas de flujo.

Más recientemente apareció la necesidad de distribuir aún más el plano de control, eliminando el problema que suponía el controlador centralizado. En este sentido, ONOS

[65] jugó un papel importante, ofreciendo la propia distribución de los controladores SDN.

2.2.3. Redes de próxima generación y gestión de la movilidad

En los últimos años, el 3GPP también ha realizado esfuerzos importantes relacionados con la separación del plano de datos y el plano de control, en el EPC (*Evolved Packet Core*) de 4G y que ha mantenido en el 5GC (*5G Core Network*) de 5G. La separación completa entre las funciones del plano de datos y las funciones del plano de control CUPS en la arquitectura del EPC evolucionado en 4G se introdujo por primera vez en la versión 14 del estándar [66], permitiendo implementaciones más flexibles, eficientes y de mayor rendimiento. Además, con la ventaja de que el procesamiento del tráfico de datos y la agregación del mismo se realicen más cerca del borde de la red, aumentando así el rendimiento global de la misma.

El concepto CUPS se ha mantenido en el tiempo y también se ha introducido en las arquitecturas 5G [42], siendo considerado en la actualidad un aspecto clave en el diseño de redes móviles 5G [47]. En este contexto, las funciones del plano de control se pueden desplegar con *software* en una plataforma *cloud*, facilitando su escalado elástico y automático. Además, las funciones del plano de datos se pueden desplegar en dispositivos de red simplificados y de alta velocidad, optimizados para el reenvío del tráfico de datos. Esta separación también permite el uso eficiente de un plano de datos común y facilita el aprovisionamiento de servicios mediante la virtualización de funciones de red (NFV) en entornos SDN, sin estar limitado únicamente a estos entornos.

Por otra parte, el IETF, a través del grupo de trabajo de DMM², está también utilizando tecnologías y mecanismos de red emergentes para rediseñar los protocolos DMM, con el objetivo de distribuir los planos de control y de datos, considerándolo un aspecto clave [8]. En la siguiente sección se explicará de manera detallada la evolución que se ha producido desde los enfoques de movilidad centralizada a los enfoques de movilidad distribuida.

2.3. Soporte a la gestión de la movilidad

Debido a los cambios e innovaciones que se han producido en los entornos de redes móviles de próxima generación, se ha incrementado la necesidad de dar

²DMM Working Group: <https://datatracker.ietf.org/wg/dmm/>

soporte a la movilidad, cuyo principal objetivo es la provisión de conectividad ubicua independientemente del lugar, instante, aplicación utilizada o tecnología de acceso empleada.

Además, a todo lo anterior se suma que el soporte a la gestión de la movilidad ha sido considerado por 5G PPP como uno de los aspectos clave de las redes de próxima generación, tal y como se ha comentado en la Sección 2.1.

Para proporcionar soporte de movilidad transparente a nivel de red se utilizan mecanismos o protocolos de gestión de la movilidad, que tienen como principal objetivo mantener abiertas (en funcionamiento) las comunicaciones cuando un dispositivo se mueve a otras redes diferentes (proceso de *handover*). A tal efecto, el IETF ha estandarizado numerosos protocolos para la gestión de la movilidad IP, que pueden ser categorizados en enfoques de movilidad centralizados y enfoques de movilidad distribuidos.

Las siguientes subsecciones muestran en detalle los enfoques más representativos utilizados hasta el momento por la comunidad científica.

2.3.1. Gestión de la movilidad centralizada

Los protocolos de gestión de la movilidad centralizada CMM [45] basan sus operaciones en la existencia de una entidad central encargada de gestionar los procesos involucrados en el soporte a la movilidad. Esta entidad central ancla la dirección IPv6 utilizada por el nodo móvil y se encarga de coordinar la gestión de la movilidad. En definitiva, el punto de anclaje central es el encargado de mantener la ubicación de los nodos móviles y redirigir el tráfico hacia su ubicación actual. Los protocolos más representativos estandarizados por el IETF han sido Mobile IPv6 (MIPv6) [67] y Proxy Mobile IPv6 (PMIPv6) [68].

Mobile IPv6, que se trata de un protocolo que ofrece soporte a la movilidad basado en el *host*³, presenta una entidad llamada *Home Agent* (HA) que funciona como ancla de movilidad. Esta entidad se ubica en la *home network* (HN) o red origen del nodo móvil (MN) y ancla a ella la dirección IP permanente o *home address* (HoA) usada por el dispositivo. Mientras el MN permanece en el dominio puede recibir paquetes destinados a su HoA que serán reenviados mediante mecanismos de encaminamiento convencionales.

³El nodo móvil tiene que tener capacidad para la detección del movimiento y poder llevar a cabo operaciones de señalización.

Cuando el MN sale de su HN obtiene una dirección IP (CoA) en la red visitada e informa al HA de su nueva ubicación a través de un mensaje *Binding Update* (BU). El HA responderá con un mensaje *Binding Acknowledgement* (BA) y establecerá un túnel bidireccional IP sobre IP con el MN; de manera que el HA capturará todo el tráfico IP dirigido al MN y lo redireccionará hacia él a través del túnel establecido entre las dos entidades.

En la Figura 2.4 se puede observar la gestión centralizada de la movilidad usando MIPv6.

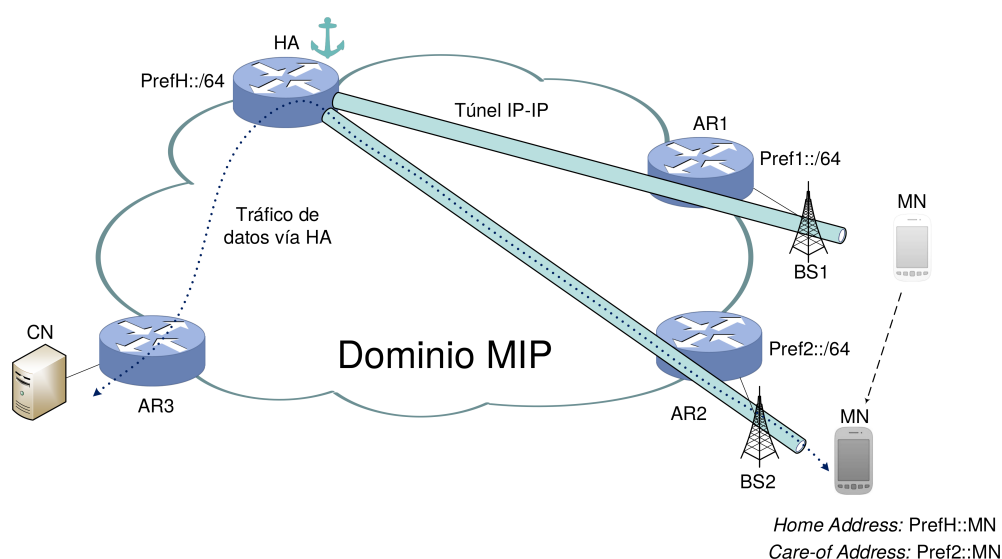


Figura 2.4: Visión general de la arquitectura de MIPv6

Por otra parte, Proxy Mobile IPv6 (PMIPv6) ofrece soporte de movilidad basado en la red⁴, introduciendo la entidad *Local Mobility Anchor* (LMA) que es la que actúa como ancla de movilidad para las sesiones abiertas de cada MN. Este protocolo reubica ciertas funcionalidades desde el MN hacia la red, concretamente la detección del movimiento y las operaciones de señalización, que serán llevadas a cabo por una nueva entidad funcional llamada *Mobile Access Gateway* (MAG). Esta entidad se suele utilizar en el borde de la red de acceso, ejerciendo de *router* de acceso (AR) para el MN, detectando los movimientos del mismo e informando al LMA de tales sucesos.

Cuando un MN realiza su movimiento por el dominio y cambia de MAG, este envía un mensaje *Proxy Binding Update* PBU al LMA para establecer un túnel IP sobre IP bidireccional entre el MAG y el LMA. Además, el LMA contesta al MAG con un mensaje de *Proxy Binding Acknowledgement* (PBA). En definitiva, se establecerán

⁴Se proporciona a los nodos móviles soporte para la movilidad sin necesidad de que intervengan de manera activa en la señalización. Por tanto, a diferencia de MIPv6, con el protocolo PMIPv6 no hay necesidad de que los nodos móviles tengan capacidades adicionales.

túneles entre el LMA y cada MAG por el que vaya pasando el MN; y gracias a ello, el LMA será capaz de capturar todo el tráfico IP dirigido al MN, pudiendo redireccionarlo correctamente hacia él mismo. Cada MN tiene configurada una *home address* con un prefijo de red gestionado por el LMA. En la Figura 2.5 se puede observar la gestión centralizada de la movilidad usando el protocolo PMIPv6.

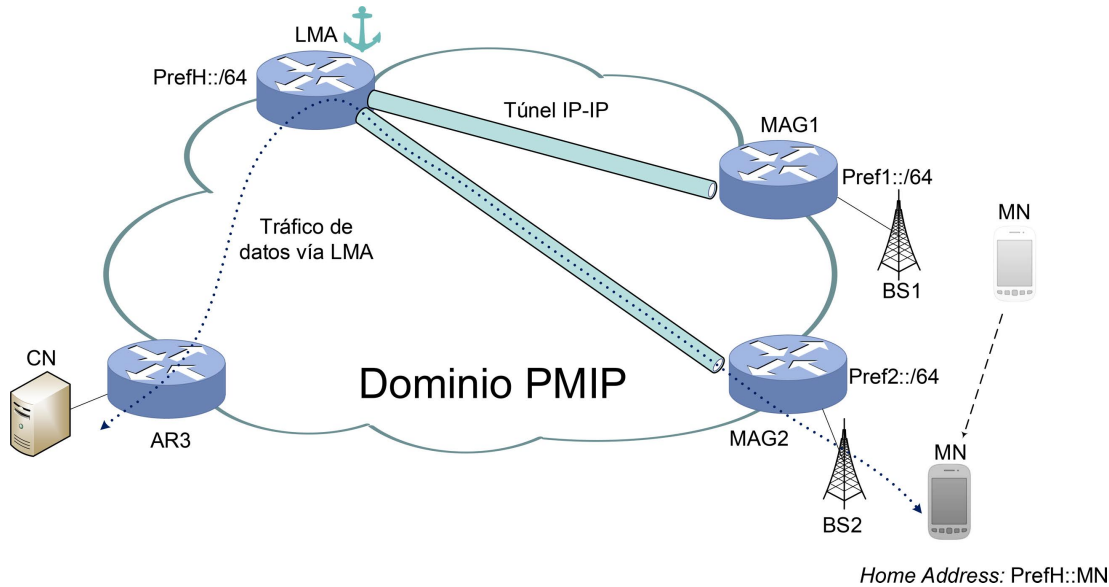


Figura 2.5: Visión general de la arquitectura de PMIPv6

Los esquemas definidos anteriormente están basados en un agente de movilidad centralizado, en el núcleo de la red, que será el encargado de capturar todo el tráfico IP dirigido hacia cada MN. La Tabla 2.1 muestra la equivalencia entre los principales roles de movilidad y cada entidad lógica, dependiendo del protocolo de movilidad utilizado.

Tabla 2.1: Equivalencia entre los principales roles de movilidad y entidades lógicas

	Mobile IPv6	Proxy Mobile IPv6
Ancla de movilidad	HA	LMA
Agente de señalización	MN	MAG

2.3.2. Gestión de la movilidad distribuida

En la actualidad, la mayoría de las arquitecturas de red desplegadas cuentan con un número limitado de agentes de movilidad centralizados que gestionan el tráfico de millones de usuarios móviles, mostrando ineficiencias cuando se maneja un gran volumen de tráfico de datos móviles. Todos estos problemas y limitaciones que presentan los enfoques centralizados han sido identificados en numerosos trabajos de investigación [9, 69, 70] y, a continuación, se resumen los más importantes.

- Largas (no óptimas) rutas de encaminamiento. Debido a que la dirección permanente utilizada por el MN se asocia a la red origen, el tráfico siempre pasará por el ancla central (HA en MIPv6 y LMA en PMIPv6), estableciéndose rutas que son, en general, más largas que la ruta directa entre el MN y el *Correspondent Node* (CN)⁵. Esto se agrava con la tendencia actual en la que los proveedores de contenidos se sitúan en el borde de la red de acceso, lo más cerca posible a los usuarios, como pasa por ejemplo con la implementación de *Content Delivery Network* o *Mobile Edge Computing*. Con enfoques CMM, el tráfico tiene que pasar primero por la red origen, añadiendo en ocasiones retardos innecesarios y desperdiciando recursos del operador.
- Problemas de escalabilidad. La escalabilidad se define como la capacidad de manejar una carga de tráfico creciente específicamente en el plano de control [71, 72]. Con la explosión del tráfico de datos móviles, el agente de movilidad de las soluciones CMM gestionará un volumen agregado de tráfico cada vez mayor, acrecentando los problemas de escalabilidad. Por tanto, los agentes de movilidad centralizados necesitan tener suficiente capacidad de procesamiento, para gestionar de manera eficiente el tráfico de control requerido para mantener el contexto de movilidad de los diferentes usuarios móviles.
- Fiabilidad. Las soluciones centralizadas son propensas a problemas de fiabilidad, debido a la entidad central, que es el único punto potencial de fallos.

Además de los principales problemas mencionados anteriormente, se pueden identificar otros, no menos importantes, como: la sobrecarga de señalización, despliegues de red más complejos y falta de granularidad en el servicio de gestión de la movilidad.

Debido a estas limitaciones, los enfoques centralizados no eran eficientes cuando se demandaba un gran volumen de tráfico de datos con requisitos de QoS y QoE. Esto motivó que las redes móviles evolucionaran hacia una arquitectura más plana⁶, apareciendo el paradigma DMM, y proporcionando así los medios eficientes y necesarios para manejar el tráfico de datos móviles. En definitiva, DMM desarrolló un nuevo concepto para gestionar la movilidad, distribuyendo las funciones de movilidad del plano de datos y del plano de control entre diferentes entidades de la red de acceso [73].

De manera similar a lo que ocurre con las soluciones CMM, dependiendo del rol del nodo móvil en el proceso de *handover*, los protocolos de gestión de movilidad distribuida pueden clasificarse principalmente en dos categorías: los que requieren una participación

⁵Es el nodo con el que se comunica e intercambia información el MN.

⁶Los agentes de movilidad son trasladados al borde de la red de acceso.

activa del MN (enfoques basados en el *host*) y los que no (enfoques basados en la red) [74]. A continuación, se describen las principales soluciones DMM basadas en el *host* y basadas en la red.

Enfoque DMM basado en el *host*

La base fundamental de este enfoque DMM (*Host-Based DMM*, *HB-DMM*) es utilizar protocolos de soporte a la movilidad basados en el *host*. La propuesta más representativa de este tipo de solución DMM está basada en Mobile IPv6 y se explica con detalle en [75, 76, 77]. HB-DMM extiende la señalización de movilidad y reutiliza muchos conceptos como la capacidad del MN para gestionar mensajes de señalización, la caché de vínculos (*Binding Cache*) en el ancla de movilidad o la tunelización del tráfico de datos. Sin embargo, el enfoque HB-DMM soporta la movilidad de una manera distribuida [75], con el objetivo de mejorar el rendimiento de la gestión de la movilidad, introduciendo agentes de movilidad distribuidos llamados *Access Mobility Anchor* (AMA). Las funciones llevadas a cabo en el HA de MIPv6 ahora son ejecutadas en los nuevos agentes de movilidad (AMA) distribuidos por la red de acceso.

Los AMA asignan prefijos de red a los nodos móviles, manteniendo la caché de vínculos actualizada, utilizando mensajes de señalización. Este enfoque distingue dos tipos de AMA:

- *Serving AMA*. Es el AMA al que está actualmente conectado y asociado el MN.
- *Origin AMA*. Es el AMA donde el MN ha configurado su dirección IP inicialmente, la cual le sirve para establecer sesiones de comunicación.

Al principio, el nodo móvil configura su dirección IP en función del prefijo de red proporcionado por el AMA. A continuación, registra la dirección configurada en el AMA enviando un mensaje BU. Cuando un MN se mueve a una red adyacente, el MN configura una nueva dirección basada en el prefijo de red obtenido del nuevo *serving AMA* de la nueva red de acceso, mientras mantiene la dirección anterior que le proporcionó el *origin AMA*. Cuando el MN se registra enviando un mensaje BU en la nueva red de acceso, registra no solo la nueva dirección que ha sido recientemente configurada, sino que también informa al *serving AMA* de la dirección anterior. En este momento, el propio *serving AMA* solicita al *origin AMA* un cambio de contexto de movilidad. Para ello utiliza nuevos mensajes de señalización: *Access Binding Update* (ABU) y *Access Binding Acknowledgement* (ABA). Como resultado de este intercambio de mensajes se crea un túnel bidireccional entre AMAs. En la Figura 2.6 se describe la arquitectura y operación de la solución HB-DMM.

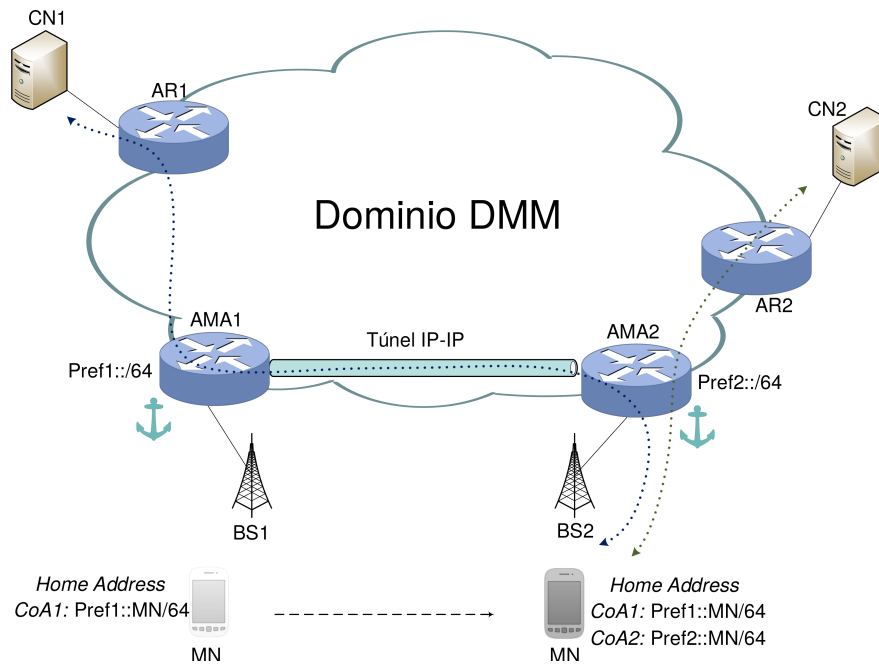


Figura 2.6: Visión general de la arquitectura de HB-DMM

Enfoque DMM basado en la red

En este tipo de enfoque basado en la red (*Network-Based DMM*, NB-DMM) [78], a diferencia de HB-DMM, el MN se encuentra exento de participar en la señalización relativa a la gestión de su propio contexto de movilidad, al igual que ocurre con PMIPv6. De esta manera, no es necesario la actualización del *software* de red para el soporte a la movilidad en el MN, ya que los agentes de movilidad distribuida llevarán a cabo las funciones de señalización requeridas. NB-DMM es una de las primeras propuestas diseñadas por el grupo de trabajo de DMM⁷.

En las arquitecturas NB-DMM, se define un nuevo tipo de entidad llamada *Mobility Access Router* (MAR) que se ubica en el borde de la red de acceso con el objetivo de obtener una red más plana separando los planos de datos y de control. Estos nuevos agentes de movilidad, además de proporcionar las funciones requeridas para anclar las sesiones de los nodos móviles a él, proporciona funciones para la actualización del contexto de movilidad de todos los nodos móviles que él mismo gestiona. El MAR detecta las conexiones de cada MN a la red de acceso y le proporciona un *Home Network Prefix* (HNP). Además, cuenta con una caché de vínculos local para almacenar la información de los nodos móviles que él mismo ha registrado en el dominio.

⁷IETF DMM Working Group: <https://datatracker.ietf.org/wg/dmm/>

Por otro lado, para mantener una tabla de correspondencias general, se define otra entidad llamada *Context Mobility Database* (CMD). Se trata de una base de datos que proporciona el contexto de movilidad de todos los nodos móviles registrados en el dominio DMM.

Cuando se produce un *handover* antes del final de la sesión, el tráfico de datos se envía a través de un túnel establecido entre el MAR actual y el MAR donde fue anclada esta sesión inicialmente. En un *handover*, el nuevo MAR recupera del CMD, a través de mensajes PBU y PBA, las direcciones IP de los MAR que actúan como ancla para las sesiones en curso del MN. Posteriormente, se procede a la actualización de la ubicación del MN en cada uno de los MAR que actúan como ancla para alguna sesión. Este proceso se lleva a cabo a través de mensajes de señalización PBU y PBA.

Adicionalmente, en las arquitecturas NB-DMM, el plano de control puede ser implementado de diferentes maneras y está sujeto a los despliegues del operador de red. Por lo tanto, de acuerdo al nivel de distribución del plano de control [69], las soluciones NB-DMM pueden ser categorizadas en:

- Soluciones parcialmente distribuidas: el plano de datos es completamente distribuido entre las distintas anclas de movilidad, mientras el plano de control es centralizado. La arquitectura y operación de este tipo de solución se muestra en la Figura 2.7.

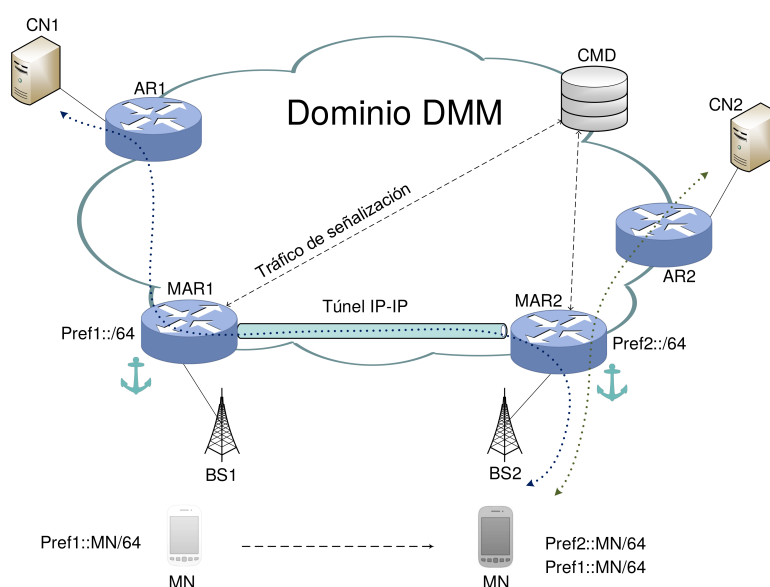


Figura 2.7: Visión general de la arquitectura de NB-DMM parcialmente distribuida

- Soluciones totalmente distribuidas: tanto el plano de datos como el plano de control son completamente distribuidos, tal y como muestra la Figura 2.8. No existe ninguna entidad de control centralizada.

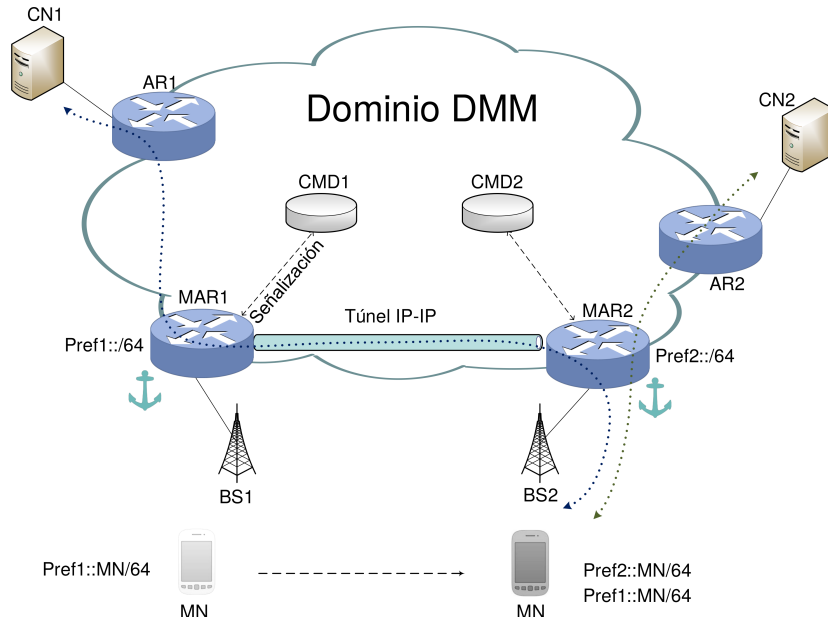


Figura 2.8: Visión general de la arquitectura de NB-DMM totalmente distribuida

Por lo tanto, la principal diferencia entre los dos diseños es si los planos de datos y control están acoplados de manera flexible (parcialmente distribuida) o estrechamente acoplados (totalmente distribuida) [8, 79].

2.3.3. Retos de la gestión de la movilidad

En estos últimos años se están discutiendo nuevos enfoques DMM diseñando un nuevo plano de control basado en SDN, aprovechando al máximo las capacidades del paradigma emergente [4, 80]. Con estos mecanismos se pretende la separación completa del plano de datos y el plano de control, controlando y asignando recursos de red dinámicamente, con el objetivo de proporcionar mayor flexibilidad y escalabilidad en el nuevo ecosistema 5G [81]. Además, estos nuevos enfoques pueden ser considerados una gran oportunidad para llevar a cabo una gestión eficiente de la movilidad en redes 5G, de tal manera que el soporte a la movilidad está basado en un servicio que se desarrolla como una aplicación SDN. La Figura 2.9 muestra la visión general de este tipo de arquitecturas.

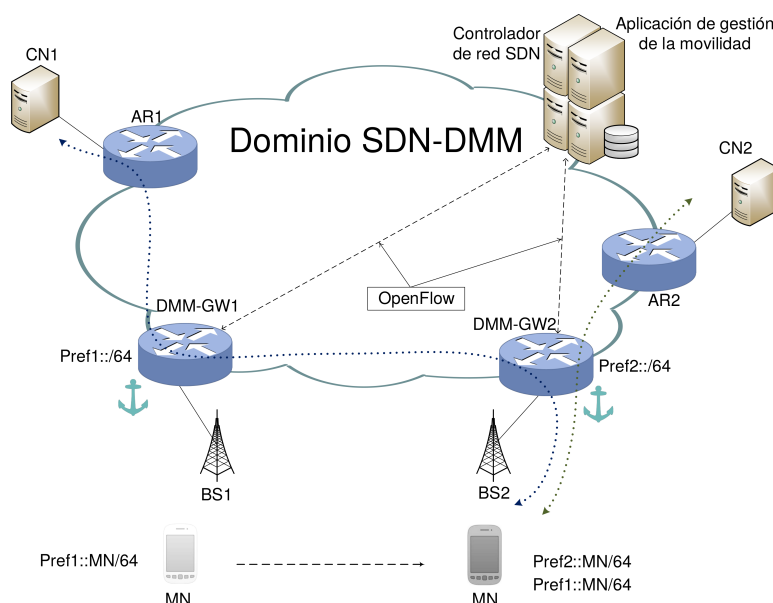


Figura 2.9: Visión general de la arquitectura de SDN-DMM

En [82], se contribuye al diseño e implementación de una solución ligera para el control de la movilidad con ONOS [65, 83]. Esta aplicación detecta la conexión de cada MN a través de mensajes de señalización *Router Solicitation* (RS) generados por el nodo IPv6, detectando si es una nueva conexión o se ha producido un *handover*. A continuación, el controlador configura un camino entre los terminales sobre SDN. En [84], se proponen dos diseños. El primero hace uso de túneles IP sobre IP en *switches* OpenFlow, introduciendo una menor sobrecarga de señalización en la modificación del flujo de datos. El segundo establece soporte para el encaminamiento óptimo, sin hacer uso de tunelización IP. Además, en [80], los autores proponen una solución que evita los túneles IP sobre IP, añadiendo la capacidad de gestión dinámica de flujos proporcionada por SDN.

2.3.4. Comparativa entre protocolos de gestión de la movilidad

En esta sección se lleva a cabo una comparativa entre los enfoques de gestión de movilidad centralizada y distribuida, analizando los principales aspectos de cada propuesta, para que se puedan establecer convenientemente diferencias y similitudes entre ellas. La Tabla 2.2 muestra dicha comparativa.

Tabla 2.2: Comparativa entre protocolos de gestión de movilidad centralizada y distribuida

	MIPv6	PMIPv6	HB-DMM	NB-DMM	SDN-DMM
Tipo de gestión de movilidad	Basada en el <i>host</i>	Basada en la red	Basada en el <i>host</i>	Basada en la red	SDN
Alcance de movilidad	Global	Local	Local	Local	Local
Infraestructura de movilidad	HA	LMA y MAG	AMA	MAR y CMD	Aplicación SDN
Modificación del MN	Requerida	No requerida	Requerida	No requerida	No requerida
Direcciones del MN	<i>Home Address y Care-of Address</i>	<i>Home Address</i>	<i>Home Address y Care-of Address</i>	Dirección IP configurada en la red de acceso	Dirección IP configurada en la red de acceso
Nº de direcciones asociadas al MN	2	1	Nº de AMAs visitados	Nº de MARs visitados	Nº de DMM-GWs visitados
Tunelización	HA → MN	LMA → MAG	AMA → AMA	MAR → MAR	No hay tunelización
Nº de túneles por MN	1	1	(Nº de AMAs visitados) - 1	(Nº de MARs visitados) - 1	0

En primer lugar, en cuanto al alcance de la movilidad, solamente MIPv6 es capaz de proporcionar un soporte a la movilidad de forma global. El resto de protocolos de movilidad analizados son capaces de gestionar la movilidad localmente, en un dominio de red específico. Otro aspecto interesante es que los protocolos basados en el *host* requieren modificaciones de la pila de protocolos en los nodos móviles, con la complejidad que ello introduce. Además, el tipo y el número de direcciones asociadas al MN varían, dependiendo de cada propuesta en cuestión. Por último, con respecto al proceso de tunelización, hay que tener en cuenta que las propuestas basadas en SDN no necesitan este tipo de mecanismos para llevar a cabo la gestión de la movilidad correctamente.

2.3.5. Principales métricas para el análisis de los protocolos de gestión de la movilidad

Para concluir este capítulo, se analizan los trabajos de investigación relacionados con la gestión de la movilidad que han sido publicados en los últimos años, identificando las métricas que nos permitan evaluar el rendimiento de las contribuciones presentadas en esta Tesis. En general, la evaluación del rendimiento de los enfoques de gestión de la movilidad basados en IP se centran principalmente en la simulación y evaluaciones experimentales.

La Tabla 2.3 muestra un resumen de los trabajos más relevantes en los cuales se realizan estudios de los protocolos de gestión de la movilidad. En ella se muestra el tipo de gestión de movilidad utilizada, los parámetros específicos analizados y el tipo de resultados obtenidos en cada trabajo. Es muy importante comprender los métodos existentes para analizar y extraer información útil. Tal y como se puede observar, las métricas utilizadas han sido bastante similares durante estos últimos años.

Trabajo	Gestión de la movilidad	Métricas	Resultados
Makaya et al., 2008 [85]	–CMM: Basado en el <i>host</i>	<i>Signaling Cost</i> <i>Packet Delivery Cost</i> <i>Buffer Space</i> <i>Handoff Latency</i>	Analíticos
Lee et al., 2010 [86]	–CMM: Basado en el <i>host</i> Basado en red	<i>Signaling Cost</i> <i>Packet Delivery Cost</i> <i>Tunnelling Cost</i> <i>Total Cost</i>	Analíticos
Lee et al., 2013 [87]	–CMM y DMM: Basados en el <i>host</i> Basados en la red	<i>Registration Delay</i> <i>Signaling Overhead</i> <i>Traffic Intensity</i>	Analíticos
Ali-Ahmad et al., 2014 [78]	–CMM y DMM: Basados en la red	<i>Signaling Cost</i> <i>Packet Delivery Cost</i> <i>Tunnelling Cost</i> <i>Processing Cost</i> <i>Handover Latency</i> <i>Handover Failure Probability</i> <i>Packet Loss</i>	Analíticos
Giust et al., 2014 [88]	–CMM y DMM: Basados en la red	<i>Signaling Cost</i> <i>Packet Delivery Cost</i> <i>Handover Latency</i> <i>Packet Loss</i>	Analíticos y Experimentales
Vasu et al., 2014 [89]	–CMM y DMM: Basados en el <i>host</i> Basados en la red	<i>Signaling Cost</i> <i>Handover Latency</i> <i>Packet Loss</i>	Analíticos y Simulados
Munir et al., 2014 [90]	–CMM y DMM: Basados en la red	<i>Number of Anchors</i> <i>Number of Tunnels</i>	Analíticos y Simulados
Jeon et al., 2015 [91]	–CMM y DMM: Basados en el <i>host</i> Basados en la red	<i>Packet Delivery Cost</i> <i>Tunnelling Overhead</i> <i>Throughput</i>	Analíticos
Park et al., 2016 [92]	–CMM: Basado en la red –SDN	<i>Signaling Cost</i> <i>Packet Delivery Cost</i>	Analíticos
Ernest et al., 2016 [93]	–DMM: Basado en la red	<i>Signaling Overhead</i> <i>Packet Delivery Cost</i> <i>End-to-End Delay</i> <i>Binding Update Cost</i>	Analíticos y Simulados

Continúa en la siguiente página.

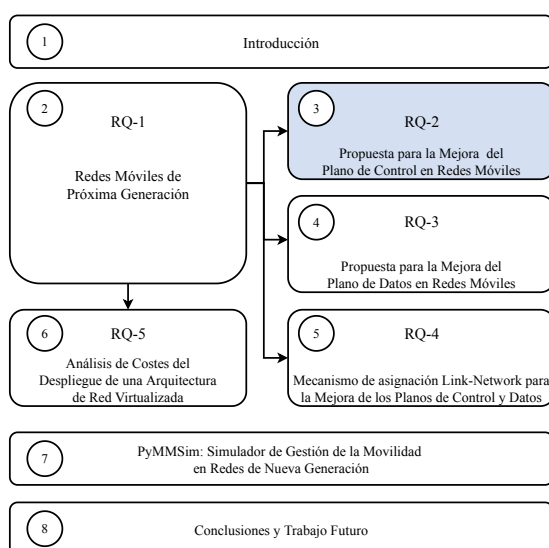
Trabajo	Gestión de la movilidad	Métricas	Resultados
Raza et al., 2016 [94]	–CMM: Basado en la red –SDN	<i>Handover Delay</i> <i>Packet Loss</i> <i>Goodput</i>	Analíticos y Experimentales
Sánchez et al., 2016 [95]	–SDN	<i>Handover Delay</i>	Experimentales
Nguyen et al., 2016 [96]	–DMM: Basado en la red –SDN	<i>Signaling Cost</i> <i>Handover Latency</i> <i>End-to-End Delay</i>	Analíticos y Experimentales
Murtadha et al., 2017 [97]	–CMM y DMM: Basados en la red	<i>End-to-End Delay</i> <i>Packet Loss</i> <i>Handover Latency</i>	Analíticos y Simulados
Cominardi et al., 2017 [98]	–DMM: Basado en la red –SDN	<i>Signaling Cost</i> <i>Handover Latency</i> <i>Flow Recovery Time</i>	Analíticos y Experimentales
Carmona et al., 2018 [46]	–CMM y DMM: Basados en el <i>host</i> Basados en la red –Hybrid DMM	<i>Signaling Cost</i> <i>Packet Delivery Cost</i> <i>Tunnelling Cost</i> <i>Mobility Anchor Load</i>	Analíticos y Simulados
Huang et al., 2018 [99]	–DMM: Basado en la red	<i>Signaling Cost</i> <i>Handover Latency</i> <i>Network Load</i> <i>Packet Loss</i> <i>Handover Block Probability</i>	Analíticos
Aman et al., 2019 [100]	–CMM y DMM: Basados en el <i>host</i> Basados en la red	<i>Handover Latency</i> <i>Packet Loss</i> <i>Handover Block Probability</i>	Analíticos y Simulados
Fafolahan et al., 2019 [101]	– DMM: Basado en la red	<i>Handover Latency</i> <i>Packet Loss</i> <i>Signaling Cost</i> <i>Packet Delivery Cost</i>	Analíticos y Simulados
Guan et al., 2019 [102]	– CMM: Basado en la red	<i>Handover Latency</i> <i>Signaling Cost</i>	Analíticos y Simulados
Calle et al., 2020 [103]	–DMM: Basado en la red	<i>Signaling Cost</i> <i>Processing Cost</i> <i>Handover Latency</i> <i>Packet Loss</i>	Analíticos y Experimentales
Calle et al., 2020 [80]	–DMM: Basado en la red –SDN	<i>Signaling Cost</i> <i>Packet Delivery Cost</i> <i>Handover Latency</i>	Analíticos y Experimentales
Balfaqih et al., 2020 [104]	–CMM y DMM: Basado en la red	<i>Handover Latency</i> <i>Handover Failure Probability</i> <i>Session Recovery</i> <i>Packet Loss</i> <i>Signaling Cost</i>	Analíticos

Tabla 2.3: Resumen de trabajos de análisis y evaluación de gestión de la movilidad

Capítulo 3

Propuesta para la Mejora del Plano de Control en Redes Móviles

En los últimos años, el tráfico global de datos móviles ha crecido exponencialmente provocando, a su vez, un incremento del tráfico de control o señalización [18, 19]. Por ello, en este capítulo se propone un nuevo mecanismo de gestión de la movilidad distribuida, llamado *Tunneling Extension to Distributed Mobility Management* (TE-DMM), que permite llevar a cabo una nueva gestión del plano de control en arquitecturas distribuidas, consiguiendo una disminución importante del tráfico de control.



El rendimiento de la propuesta TE-DMM ha sido evaluado de manera analítica y experimental con respecto a la solución más representativa de gestión de la movilidad distribuida basada en la red (NB-DMM), cuya operación se describe detalladamente en la Sección 2.3.2, y la visión general de la arquitectura se muestra en la Figura 3.1. En la evaluación del rendimiento y la eficiencia se presta especial atención a los costes de señalización y a los costes de procesamiento; además de llevar a cabo un análisis y evaluación experimental de la latencia de *handover* y de la pérdida de paquetes.

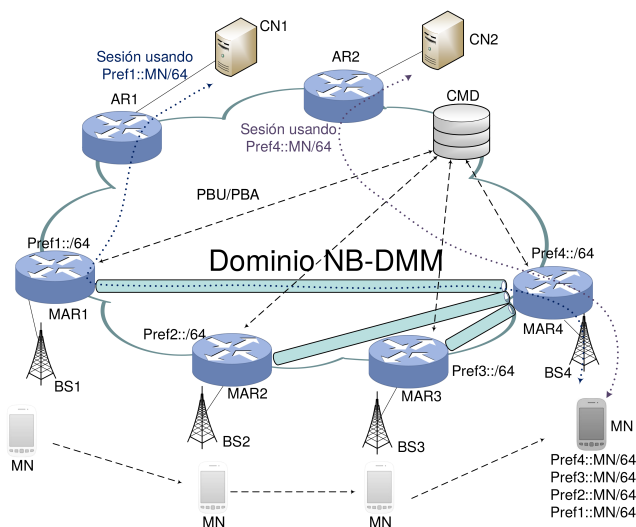


Figura 3.1: Arquitectura de NB-DMM

En definitiva, el capítulo se organiza tal y como se presenta a continuación:

- La Sección 3.1 describe de manera detallada la arquitectura propuesta para la gestión eficiente del plano de control en redes móviles, llamada TE-DMM.
- En la Sección 3.2 se desarrolla un modelo analítico que permite evaluar el rendimiento de la propuesta TE-DMM con respecto a la solución más representativa de gestión de la movilidad distribuida (NB-DMM). Además, se definen las métricas utilizadas para la evaluación de la propuesta.
- En la Sección 3.3 se presentan y comparan analíticamente los resultados obtenidos para ambas soluciones (NB-DMM y TE-DMM).
- La Sección 3.4 describe la implementación desarrollada de la propuesta TE-DMM y presenta una evaluación experimental comparativa de ambas soluciones sobre un *testbed* de movilidad, usando implementaciones reales en ambos casos.
- La Sección 3.5 muestra las conclusiones que han sido identificadas a partir del análisis realizado y la evaluación llevada a cabo.

3.1. Arquitectura propuesta TE-DMM

Los enfoques DMM basados en la red (NB-DMM), también conocidos como *PMIPv6-based DMM*, introducen principalmente dos entidades funcionales llamadas MAR (*Mobility Access Router*) y CMD (*Centralized Mobility Database*) [8, 78], tal y como se explicó en la Sección 2.3.2. Cada MAR actúa como agente de movilidad situado, generalmente, en el borde de la red de acceso, proporcionando funciones de localización para el MN. Además, estos agentes de movilidad se encargan de mantener las sesiones activas cuando los nodos móviles realizan su movimiento entre diferentes redes de acceso, proporcionando la detección y negociación del *handover* a nivel IP. Por otra parte, el CMD constituye una base de datos que contiene información de los nodos móviles registrados en el dominio de movilidad en cuestión.

El principal objetivo de TE-DMM consiste en subsanar los problemas que aparecen generalmente en los enfoques de gestión de la movilidad distribuida, desde el punto de vista del plano de control, reduciendo así la carga de señalización y, por consiguiente, la latencia de *handover* y la pérdida de paquetes [105].

Para ello, en TE-DMM hemos incorporado una nueva entidad llamada T-MAR (*Tunneling-Mobility Access Router*), que realiza las mismas funciones del agente MAR y, además, modifica la gestión de la tunelización, tal y como se puede observar en la Figura 3.2. Con esta nueva gestión del plano de control se consiguen importantes beneficios en él, reduciendo el tráfico de señalización en gran medida.

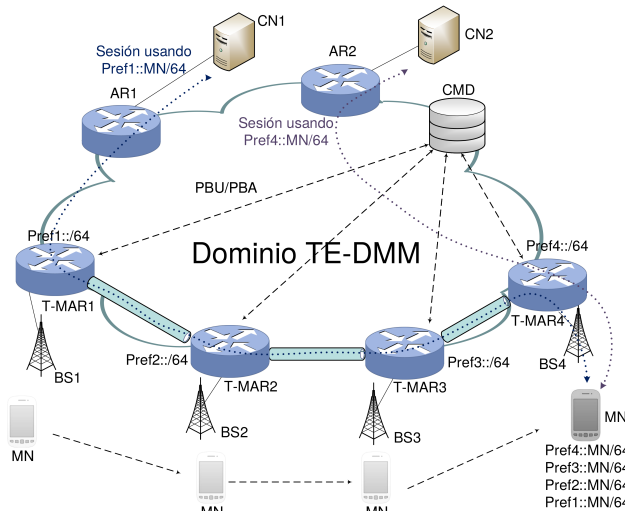


Figura 3.2: Visión general de la arquitectura de red propuesta: TE-DMM

En la solución NB-DMM (Figura 3.3), una vez que el MN realiza el movimiento, envía un mensaje *Router Solicitation* (RS) al MAR de acceso ①, el cual se encarga de notificar al CMD de tal hecho mediante un mensaje de *Proxy Binding Update* (PBU) ②. Cuando llega este mensaje, el CMD responde con un mensaje *Proxy Binding Acknowledgement* (PBA) ③. El MAR recupera la dirección IP del agente que actúa como ancla de movilidad para las sesiones del nodo móvil desde el CMD y, posteriormente, envía un mensaje *Router Advertisement* (RA) al MN con el nuevo prefijo de red ④. Luego, desde el CMD se informa a través de un mensaje PBU a todos los MAR anteriores por los que ha ido pasando el MN y actúan como ancla, para que se actualicen los túneles que se encuentran creados en ese momento ⑤. Para finalizar, dichas anclas de movilidad responden con un mensaje PBA al CMD, además de llevar a cabo la actualización de túneles ⑥. La Figura 3.3 muestra el diagrama de flujo de mensajes de la solución NB-DMM.

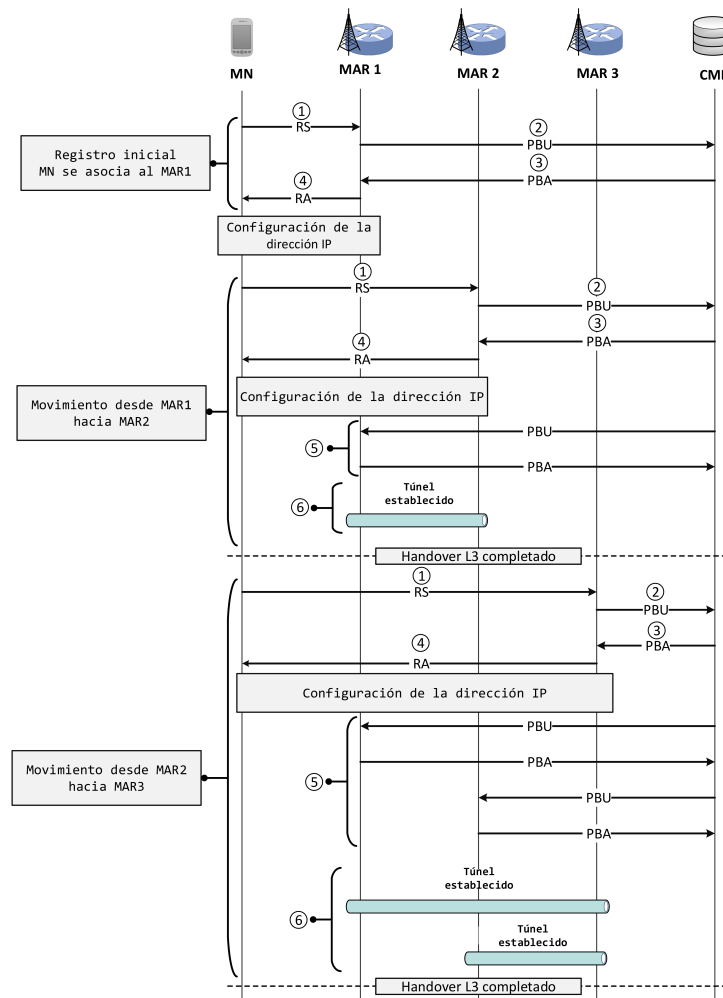


Figura 3.3: Diagrama de flujo de mensajes en NB-DMM

En la propuesta TE-DMM, la gestión del plano de control es optimizada reduciendo el número de túneles que es necesario establecer y/o eliminar durante el movimiento de un nodo móvil. La idea principal en TE-DMM es extender un único túnel desde el T-MAR anterior hasta el T-MAR en el que se encuentra el MN en un momento determinado. Estos túneles se encargan de manejar debidamente el tráfico del plano de datos. El diagrama de flujo de mensajes de TE-DMM se muestra en la Figura 3.4.

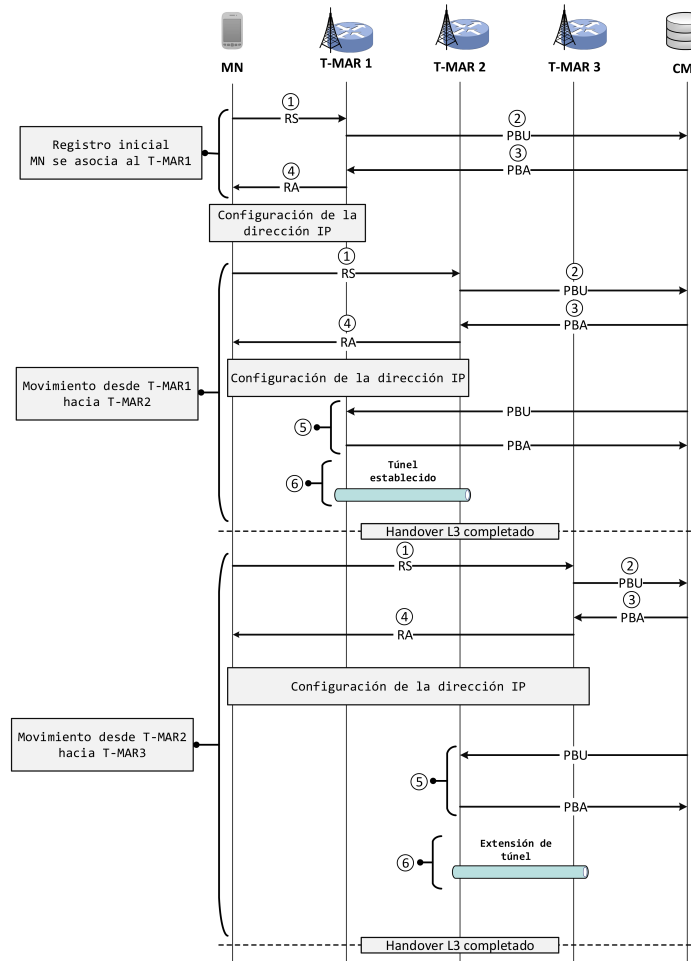


Figura 3.4: Diagrama de flujo de mensajes en TE-DMM

El funcionamiento básico de TE-DMM desde el punto de vista del plano de control es el siguiente: cuando el MN lleva a cabo un *handover*, este envía un mensaje RS al T-MAR de acceso que le está dando servicio (1), el cual informa al CMD de dicha asociación a través de un mensaje PBU (2). Cuando este mensaje es recibido por el CMD, se actualiza la información del nodo móvil y se envía un mensaje de confirmación PBA al T-MAR correspondiente (3). Además, se envía un mensaje *Router Advertisement* (RA) al MN con el nuevo prefijo de red (4). Posteriormente,

se envía un mensaje PBU al T-MAR previo ⑤, para realizar la extensión del túnel ⑥ correctamente. Como se puede comprobar, la principal diferencia con respecto a NB-DMM es que no hay que notificar a todos los agentes de movilidad previos, sino solamente al anterior, reduciendo así la carga de señalización y de tiempos asociados. Por lo tanto, en cada *handover* se establecerá solamente un túnel con el T-MAR previo.

Como se ha comentado en reiteradas ocasiones, uno de los aspectos más importantes de TE-DMM es la gestión de flujos IPv6. Por ello, para clarificar las diferencias entre la solución NB-DMM y la propuesta TE-DMM, se presenta una comparativa de la manera de proceder cuando se manejan flujos de datos IPv6.

La Figura 3.5 muestra la situación de ambas propuestas cuando un MN se conecta al dominio por primera vez. Como se puede observar, no se aprecian diferencias entre ambas soluciones en dicha situación.

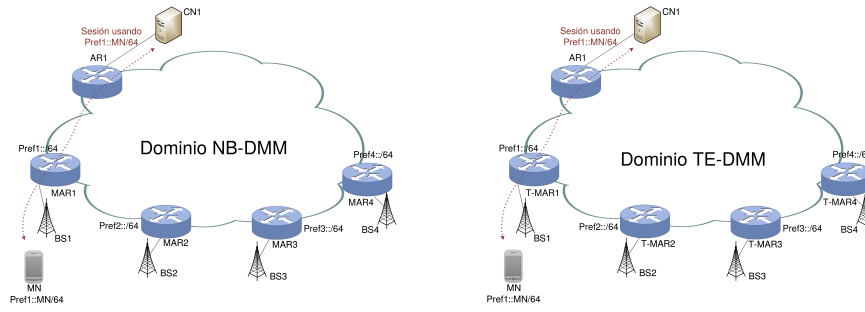


Figura 3.5: Primera fase de gestión de flujos: conexión del MN al MAR1/T-MAR1

Lo mismo ocurre cuando el MN realiza su primer movimiento, es decir, deja el MAR1/T-MAR1 y se asocia al segundo agente de movilidad (MAR2/T-MAR2). Simplemente se lleva a cabo el establecimiento de un túnel bidireccional entre ambos agentes de movilidad y se tuneliza todo el tráfico anclado al MAR1/T-MAR1 hacia el MAR2/T-MAR2, tal y como muestra la Figura 3.6.

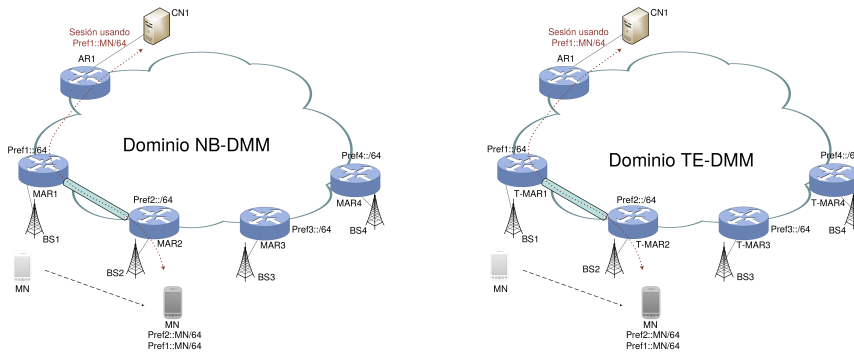


Figura 3.6: Segunda fase de gestión de flujos: movimiento del MN al MAR2/T-MAR2

A partir del primer movimiento se aprecian diferencias considerables entre NB-DMM y la propuesta TE-DMM, tal y como muestra la Figura 3.7. En la tercera fase de NB-DMM, Figura 3.7(a), cuando el MN realiza su segundo movimiento, se elimina el túnel que estaba creado y se establecen túneles nuevos desde el MAR actual hacia todos los MAR previos. Sin embargo, en la tercera fase de la propuesta TE-DMM, Figura 3.7(b), simplemente se extiende el túnel que se encontraba creado desde la segunda fase.

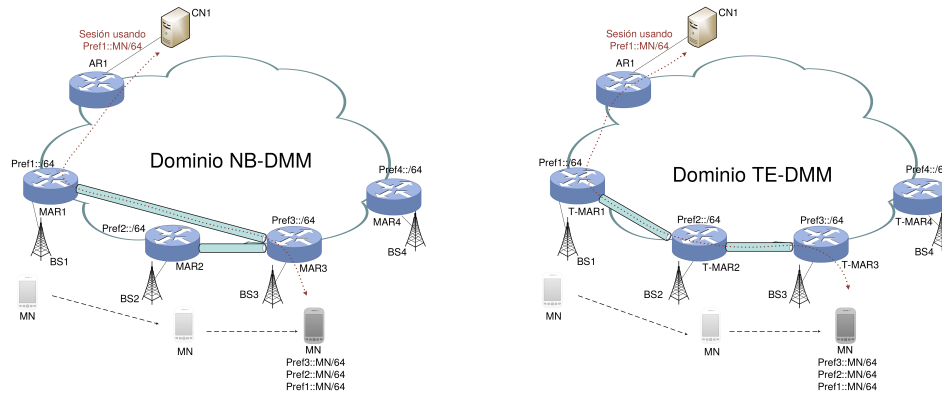


Figura 3.7: Tercera fase de gestión de flujos: movimiento del MN al MAR3/T-MAR3. Creación de nuevos túneles (a) y extensión de túnel de T-MAR2 a T-MAR3 (b).

En la cuarta fase ocurre lo mismo que en la anterior, en el caso de NB-DMM, Figura 3.8(a), se eliminan los túneles creados en la fase anterior y se establecen nuevos túneles desde el MAR actual hacia los MAR previos. Por otro lado, en la propuesta TE-DMM, Figura 3.8(b), se extiende el túnel que se encontraba creado desde la fase anterior.

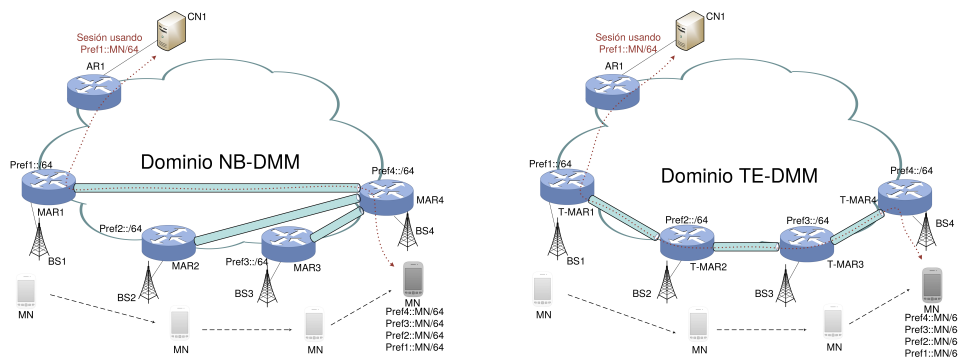


Figura 3.8: Cuarta fase de gestión de flujos: movimiento del MN al MAR4/T-MAR4. Creación de nuevos túneles (a) y extensión de túnel de T-MAR3 a T-MAR4 (b).

Si el MN siguiera realizando movimientos por la red de acceso, el comportamiento de ambas propuestas sería el mismo que en las dos últimas fases, para cada caso. En NB-DMM se eliminarían y modificarían los túneles anteriores, además de establecer

nuevos túneles desde el MAR actual hacia todos los MAR previos. Por el contrario, en la propuesta TE-DMM se extenderá el túnel creado en la segunda fase, tantas veces como movimientos realice el MN por la red de acceso.

3.2. Modelo analítico

En esta sección, se desarrolla un modelo analítico para evaluar el rendimiento de la solución de gestión de la movilidad distribuida (NB-DMM) y nuestra propuesta (TE-DMM) de forma comparativa, en términos de costes de señalización, costes de procesamiento y pérdida de paquetes.

Se considera un escenario de red en el que un MN puede participar activamente con varios CN en Internet de manera simultánea, es decir, cada MN puede tener varias sesiones activas a la vez. Se asume que el número de sesiones entrantes a un MN sigue un proceso de Poisson con una tasa media λ_S , es decir, el tiempo entre llegadas de sesiones está exponencialmente distribuido con dicha tasa. En cuanto al modelado del tráfico, se utilizará un sistema de colas M/M/ ∞ [46, 78, 106], en el que, tal y como se ha mencionado anteriormente, las sesiones entrantes siguen un proceso de Poisson con una tasa media λ_S . Además, es asumido que la duración de una sesión está exponencialmente distribuida con parámetro μ_S [107, 85]. Por lo tanto, si modelamos el escenario como un sistema bajo la distribución de probabilidad de una cola típica M/M/ ∞ , el número medio de sesiones activas (ρ_S) en un momento determinado puede calcularse como se muestra en la Ecuación 3.1.

$$\rho_S = \frac{\lambda_S}{\mu_S} \quad (3.1)$$

Por otro lado, con respecto al modelo de movilidad, se ha usado un modelo comúnmente empleado en la literatura, llamado *Fluid-Flow Model* [108, 109, 110]. Se trata de un modelo ampliamente utilizado para caracterizar el tiempo medio de estancia de un MN en una subred determinada. Con ello, se modela el movimiento del MN por la red, considerando la tasa de realización de un *handover* igual a la tasa de traspaso de una región de cobertura a otra en un área determinada. El modelo asume que la densidad de usuarios es uniforme en toda esta área, y que la dirección del movimiento está uniformemente distribuida en un rango $[0, 2\pi)$, donde v es la velocidad media de movimiento del MN; L_S y A_S son el perímetro y el área de cada región, respectivamente. De esta manera, el modelo *Fluid-Flow* establece que la tasa media de paso de una región a otra puede expresarse tal y como se muestra en la Ecuación 3.2.

$$\mu_{SR} = \frac{v \cdot L_S}{\pi \cdot A_S} \quad (3.2)$$

La Figura 3.9 muestra la distribución de celdas en un área concreta. Consideramos una red hexagonal que consiste en k anillos de celdas hexagonales. La celda 0 es considerada como la celda central; las celdas etiquetadas con 1 corresponden al primer anillo alrededor de la celda 0. Cada anillo es etiquetado de acuerdo a su distancia al centro [108, 78], de tal manera que el anillo a_k se refiere al anillo k -ésimo desde la celda central, cuya distancia al centro es igual a k . Tal y como se muestra en la Figura 3.9, el número de celdas en el anillo k -ésimo son $6k$. Por lo tanto, el número total de celdas $N(k)$ en una determinada región es definida en la Ecuación 3.3.

$$N(k) = \sum_{i=1}^k 6 \cdot i + 1 = 3 \cdot (k + 1) \cdot k + 1 \quad (3.3)$$

Consecuentemente, estableciendo el lado l de dichas celdas hexagonales, se puede determinar su perímetro L_S y el área A_S , tal y como se muestra en la Ecuación 3.4. Además, también se puede definir la tasa de traspaso entre celdas como se muestra en la Ecuación 3.5.

$$L_S = 6 \cdot l \quad A_S = \frac{3 \cdot \sqrt{3} \cdot l^2}{2} \quad (3.4)$$

$$\mu_{SR} = \frac{v \cdot L_S}{\pi \cdot A_S} = \frac{v}{l} \cdot \left(\frac{4 \cdot \sqrt{3}}{3 \cdot \pi} \right) \quad (3.5)$$

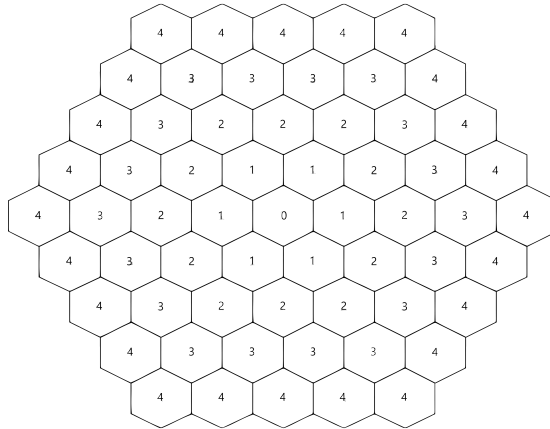


Figura 3.9: Distribución de celdas en un área de cobertura

Se considerará que todas las celdas tienen la misma área y, además, cada celda estará asociada a un *router* de acceso, constituyendo el punto de conexión entre el nivel físico y el nivel IP.

Por otro lado, el tiempo medio de estancia en una celda determinada T_{SR} para un nodo móvil puede calcularse tal y como se muestra en la Ecuación 3.6.

$$T_{SR} = \frac{1}{\mu_{SR}} \quad (3.6)$$

3.2.1. Coste de señalización

El proceso para garantizar que la sesión de movilidad esté actualizada mientras un MN se mueve entre distintas redes es una de las principales funcionalidades para cualquier protocolo de gestión de movilidad. Esta tarea requiere de mensajes de control que son enviados entre los agentes de movilidad correspondientes, llevando a cabo una gestión desde el punto de vista del plano de control que tiene asociado un coste llamado coste de señalización C_S [85, 86, 98]. Este coste depende del tamaño de los mensajes de señalización y del número de saltos en cada proceso de *handover* de nivel 3, durante el intervalo de tiempo en el que la comunicación del MN tiene que permanecer activa.

En NB-DMM, $C_S(NB - DMM)$ puede ser calculado como el coste que supone actualizar los vínculos de movilidad del MN en el CMD, además del coste de actualización de todos los MAR previos a los que se ha ido asociando el MN, ya que necesitan actualizar y crear nuevos túneles bidireccionales. Por lo tanto, el coste de señalización durante el movimiento de un MN para la solución NB-DMM se encuentra representado en la Ecuación 3.7.

$$C_S(NB - DMM) = \mu_{SR} \cdot \left(C_{CMD} + \sum_{i=1}^{n-1} (C_{MAR})_i \right) \quad (3.7)$$

donde n define el número de MAR previos que tienen que establecer un túnel con el MAR actual.

Además, el coste de actualización de un determinado agente de movilidad (MAR) es definido en la Ecuación 3.8.

$$C_{MAR} = 2 \cdot s_c \cdot h_{CMD-MAR} \quad (3.8)$$

donde s_c representa el tamaño medio de los mensajes de control y $h_{CMD-MAR}$ la distancia en número de saltos desde el CMD al MAR.

Por otro lado, en TE-DMM ($C_S(TE - DMM)$), hay que actualizar el CMD y únicamente el T-MAR previo, es decir, solamente es necesario extender el túnel desde el T-MAR previo al T-MAR actual. Por lo tanto, el coste de señalización para la propuesta TE-DMM es definido en la Ecuación 3.9.

$$C_S(TE - DMM) = \mu_{SR} \cdot (C_{CMD} + C_{TMAR}) \quad (3.9)$$

Además, el coste de actualización de un determinado agente T-MAR es definido en la Ecuación 3.10.

$$C_{TMAR} = 2 \cdot s_c \cdot h_{CMD-TMAR} \quad (3.10)$$

donde s_c es el tamaño medio de los mensajes de control y $h_{CMD-TMAR}$ es la distancia en número de saltos entre el CMD y el T-MAR en cuestión.

En ambos casos, el coste de actualizar el vínculo de movilidad en el CMD es expresado tal y como muestra la Ecuación 3.11.

$$C_{CMD} = 2 \cdot s_c \cdot h_{MAR-CMD} = 2 \cdot s_c \cdot h_{TMAR-CMD} \quad (3.11)$$

donde s_c representa el tamaño medio de los mensajes de control y h_{x-y} es la distancia en número de saltos desde el nodo x al nodo y en el dominio de movilidad.

3.2.2. Coste de procesamiento

El coste de procesamiento (C_P) representa el número de mensajes de señalización procesados por una entidad de red y por unidad de tiempo [78]. Esta métrica es relevante debido a que valores altos de ella reflejarán una mayor probabilidad de tener problemas de escalabilidad, que suponen una preocupación importante en los protocolos actuales de gestión de la movilidad. Por este motivo, se ha decidido evaluar el coste de procesamiento asociado al CMD. El interés en esta entidad es debido a que constituye un punto común que almacena todos los vínculos de movilidad para los nodos móviles de un dominio de red determinado, y puede ser visto como un potencial punto de fallo.

En definitiva, para ambas propuestas, se considera C_P como el número de mensajes de señalización procesados por unidad de tiempo en el CMD, tal y como se muestra en las Ecuaciones 3.12 y 3.13.

$$C_P(NB - DMM) = 2 \cdot n \cdot \mu_{SR} \quad (3.12)$$

$$C_P(TE - DMM) = 4 \cdot \mu_{SR} \quad (3.13)$$

3.2.3. Pérdida de paquetes

La latencia de *handover* (LH) es una métrica crítica que tiene un gran impacto en el rendimiento global de la red. Puede ser definida como el intervalo de tiempo en el cual un MN no tiene conectividad como resultado de un *handover*. Este proceso se produce cuando un MN cambia de punto de acceso produciéndose un tiempo de desconexión. LH está formada por latencia de *handover* de nivel 2 (T_{L2}), el tiempo de detección de movimiento (T_{DM}) y el tiempo necesario para actualizar las sesiones activas del MN (T_{AC}), tal y como muestra la Ecuación 3.14.

$$LH = T_{L2} + T_{DM} + T_{AC} \quad (3.14)$$

Relacionado con el proceso de *handover*, otra métrica relevante es la pérdida de paquetes durante dicho proceso. El número de paquetes perdidos durante un *handover* es directamente proporcional a la latencia de *handover* y a la tasa de llegada de paquetes ($\mu_S \cdot P_S$). Por lo tanto, la pérdida de paquetes (PL) se puede calcular como muestra la Ecuación 3.15.

$$PL = \rho_S \cdot \mu_S \cdot P_S \cdot LH \quad (3.15)$$

donde ρ_S representa el número medio de sesiones activas, μ_S es la duración de una sesión y P_S representa la media del número de paquetes de una sesión.

3.3. Evaluación de rendimiento

En esta sección se presentan y discuten los resultados numéricos obtenidos de manera analítica. NB-DMM y TE-DMM son evaluadas en términos de costes de señalización, costes de procesamiento y pérdida de paquetes. En definitiva, se estudiará el impacto de diferentes parámetros sobre los costes de movilidad.

A continuación se muestran los parámetros por defecto que constituyen la configuración básica de la topología, del modelo de movilidad y del modelo de tráfico usada en el análisis [9, 8, 46, 78, 85].

- Lado de las celdas hexagonales: $l = 500$ m.
- Velocidad del MN: $v = 20$ m/s.
- Tamaño de los mensajes de control: $s_c = 80$ bytes.
- Número de saltos entre las entidades de red: $h_{MAR-CMD} = h_{CMD-MAR} = h_{TMAR-CMD} = h_{CMD-TMAR} = 12$ saltos.
- Número de MAR/T-MAR en la topología de red: $n = 3$.
- Número medio de paquetes en una sesión: $P_S = 1000$.
- Duración de una sesión: $1/\mu_S = 20$ s.
- Número de sesiones entrantes por MN: $\lambda_S = 40/3600$ s $^{-1}$.

A partir de estos datos, en primer lugar se ha llevado a cabo un análisis del impacto de la velocidad del MN en el coste de señalización, tal y como se muestra en la Figura 3.10.

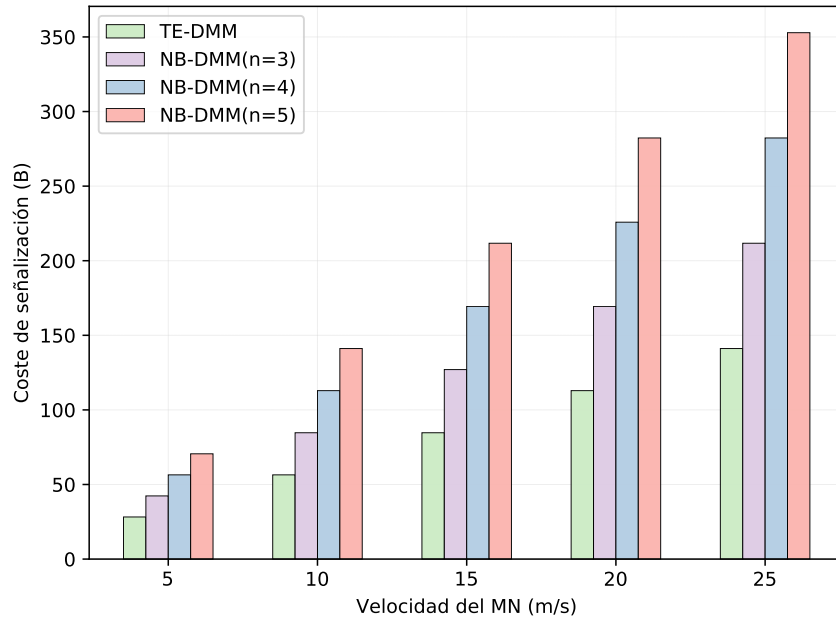


Figura 3.10: Coste de señalización dependiendo de la velocidad del MN

Sobre las mismas condiciones de red, se ha variado la velocidad del MN desde 5 a 25 m/s. Además, se ha considerado un número variable de agentes de movilidad que varía entre 3 y 5. Este efecto solamente ha sido estudiado para la solución NB-DMM, ya que en TE-DMM únicamente es necesario actualizar el T-MAR previo, por lo que el coste de señalización no depende del número de agentes de movilidad.

Por tanto, el coste de señalización es directamente proporcional a la velocidad del MN v para ambas soluciones, tal y como se muestra en la Figura 3.10, obteniéndose mejores resultados con la propuesta TE-DMM, debido a la gestión eficiente del plano de control que ella realiza.

En cuanto al coste de señalización, también se ha evaluado el efecto del perímetro de la celda de cobertura hexagonal, variando la longitud del lado de las celdas (l) entre 400 y 800 metros, como se muestra en la Figura 3.11.

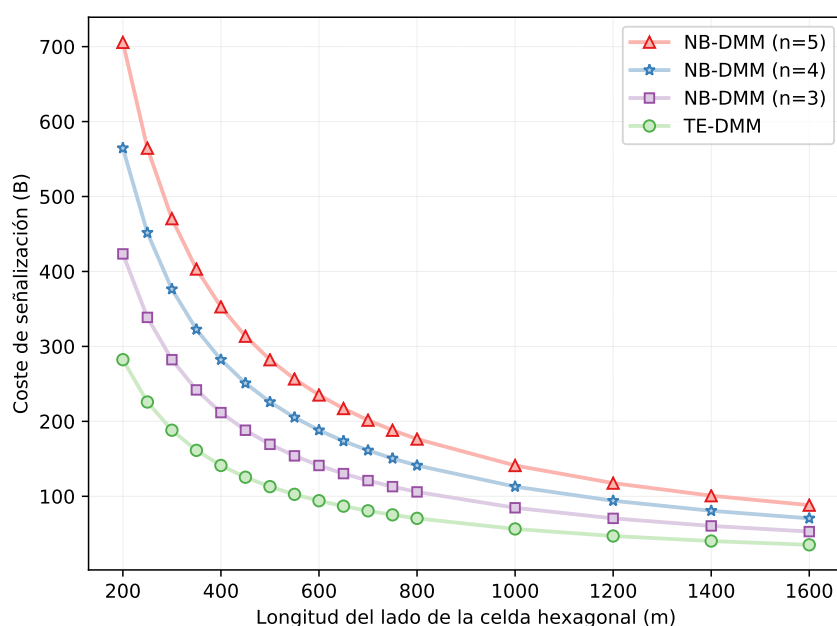


Figura 3.11: Coste de señalización dependiendo de la longitud del lado de la celda hexagonal

Para el efecto del tamaño de la celda de cobertura se han tomado los valores por defecto, excepto l . Tal y como se puede observar en la Figura 3.11, conforme aumenta la longitud del lado de la celda de cobertura hexagonal, el coste de señalización disminuye debido a que la tasa de *handover* también disminuye proporcionalmente. Esto produce un decremento de las actualizaciones de las sesiones activas sobre los agentes de movilidad. Además, se observa una mejora del coste de señalización para la propuesta TE-DMM cuando el valor del lado de la celda hexagonal incrementa. En definitiva, el coste de señalización en TE-DMM es menor que el introducido por NB-DMM.

También se ha evaluado el coste de procesamiento con respecto al número de agentes de movilidad activos para un MN. Se puede observar en la Figura 3.12 que, en NB-DMM, el coste de procesamiento es mayor conforme aumenta el número de MAR. Sin embargo, con TE-DMM el valor de dicho coste de procesamiento se mantiene constante debido a que no hay que actualizar todos los T-MAR previos, sino solamente el último que dio servicio al MN, reduciendo tiempos de procesamiento.

De esta manera, TE-DMM consigue ganancias significativas relacionadas con los costes de procesamiento y señalización, proporcionando mejoras sustanciales en el plano de control.

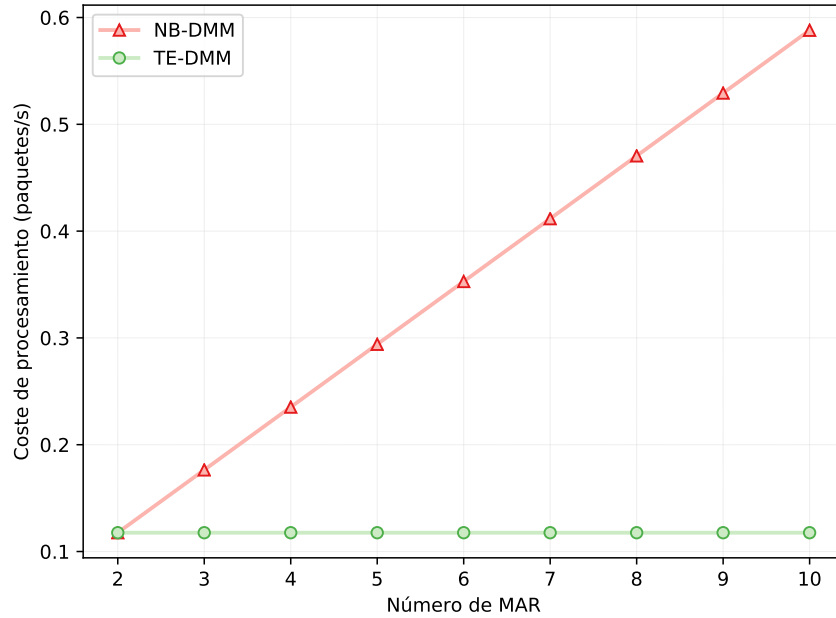


Figura 3.12: Coste de procesamiento dependiendo del número de anclas de movilidad

Por último, tal y como se muestra en la Figura 3.13, se ha llevado a cabo un análisis de la pérdida de paquetes en función del número medio de paquetes de cada sesión en un MN (ρ_S). Como se puede observar, los resultados confirman que la pérdida de paquetes incrementa con respecto al número medio de paquetes en una sesión. Estos resultados muestran una tendencia lineal para ambas soluciones, obteniendo un mejor rendimiento con la propuesta TE-DMM.

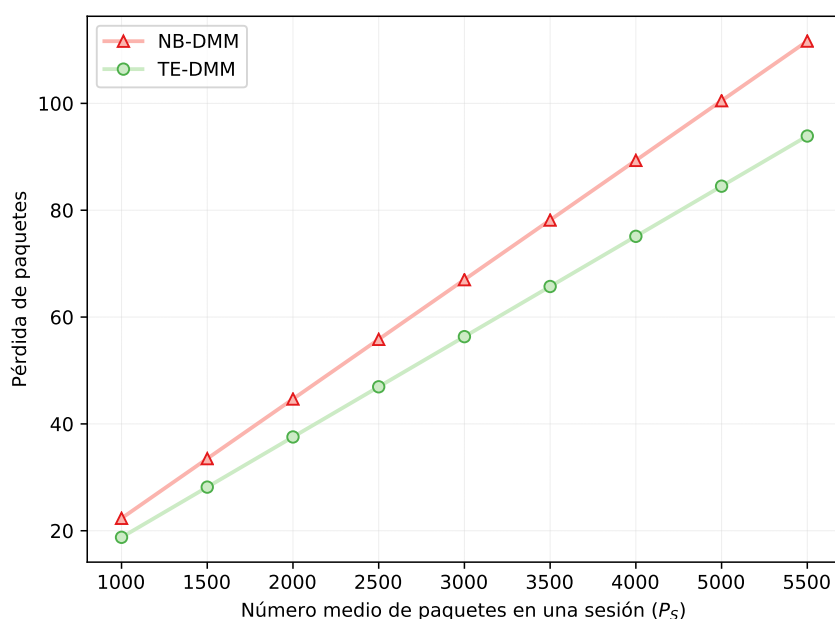


Figura 3.13: Pérdida de paquetes dependiendo del número medio de paquetes de una sesión

3.4. Evaluación experimental

Esta sección presenta una evaluación experimental comparativa utilizando implementaciones reales de NB-DMM y nuestra propuesta TE-DMM.

En el caso de NB-DMM se ha utilizado una implementación para sistemas *GNU/Linux* conocida como MAD-PMIPv6 (*Mobility Anchors Distribution for Proxy Mobile IPv6*) [111, 112], la cual está desarrollada en código C sobre el *kernel* de Linux.

Por otro lado, la propuesta TE-DMM toma como base y punto de partida el enfoque anterior MAD-PMIPv6. Sobre ella, se ha incorporado la gestión eficiente del plano de control, implementando el mecanismo de extensión de túneles en código C, y sobre el *kernel* de Linux.

En relación a la implementación de la propuesta TE-DMM de extensión de túneles, como punto de partida, se ha elaborado un diagrama para identificar los puntos claves a implementar sobre el código de la solución. En la Figura 3.14 se muestra el funcionamiento detallado de la propuesta TE-DMM, desde el punto de vista del plano de control, permitiendo clarificar el comportamiento que deben tener los agentes de movilidad (T-MAR), ya que la implementación únicamente se ejecutará sobre ellos.

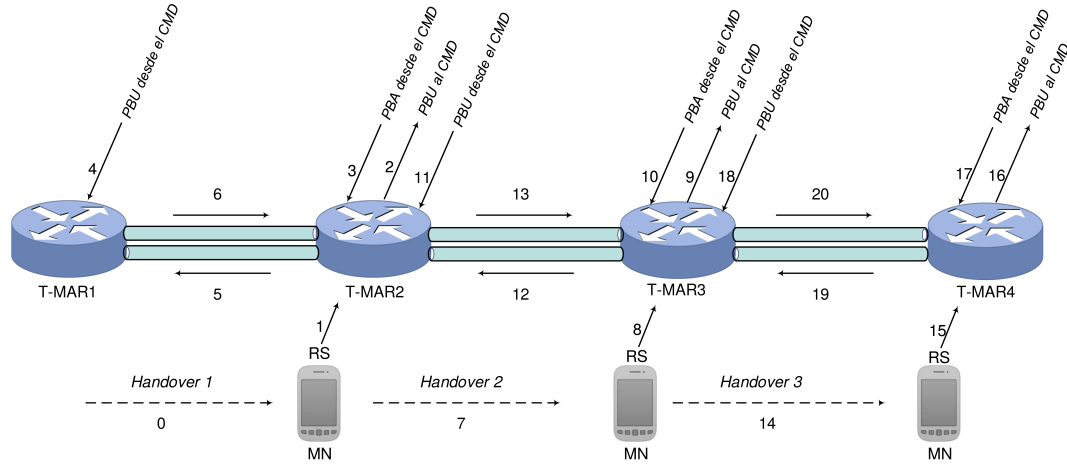


Figura 3.14: Procedimiento de extensión de túneles en la propuesta TE-DMM

Por tanto, se ha modificado el comportamiento de los T-MAR a través de la implementación de dos funciones, que se ejecutarán cuando se produzcan unos eventos específicos involucrados en la gestión de la movilidad.

Antes de explicar las funciones implementadas exhaustivamente, se define un parámetro de entrada común a ambas. Se trata de un registro, llamado *tedmm_entry*, que almacenará toda la información del nuevo vínculo de movilidad que se creará cada vez que el MN realice un proceso de *handover*. Este registro cuenta con una serie de campos, que son definidos y explicados en el Pseudocódigo 3.1.

Pseudocódigo 3.1: Registro con la información de un vínculo de movilidad

```

struct tedmm_entry{
    struct in6_addr    mn_prefix;           ▷ Prefijo de red del MN
    struct in6_addr    our_addr;            ▷ Dirección del T-MAR que ejecuta la función
    struct in6_addr    mn_hw.address;       ▷ Dirección MAC del MN
    struct in6_addr    mn_addr;            ▷ Dirección IPv6 del MN
    struct in6_addr    mn_serv_tmar_addr;   ▷ Dirección IPv6 del T-MAR actual del MN
    struct in6_addr    mn_serv_cmd_addr;    ▷ Dirección IPv6 del CMD
    struct in6_addr    tmar_hw.address;     ▷ Dirección MAC del T-MAR
    struct in6_addr    mn_prev_tmar_addr;   ▷ Dirección IPv6 del T-MAR previo del MN
    struct in6_addr    mn_link_local_addr;  ▷ Dirección IPv6 de enlace local del MN
};

```

Tal y como se ha comentado anteriormente, este registro será utilizado por las funciones implementadas para la propuesta TE-DMM, cuya operación es descrita a continuación:

- Función *“registration_serving_tmar”*: esta función será ejecutada cuando cualquier T-MAR reciba un mensaje PBA. En el Algoritmo 1 se muestra su implementación y operación. Con esta función se finaliza el registro de un MN conectado localmente, actualizando su vínculo de movilidad y respondiendo al MN con un

mensaje RA que contenga el nuevo prefijo de red ③. Además, esta función será la encargada de establecer el túnel desde el T-MAR actual hasta el T-MAR previo ① y actualizar las rutas en el agente de movilidad que ejecuta la función ②. En este momento estaría creado el túnel *uplink* hasta el T-MAR previo, pero hay que tener en cuenta que los túneles deben ser bidireccionales; por lo que es necesario ejecutar otra función que se encargue de completar el procedimiento de extensión del túnel.

Algoritmo 1 Registro de un MN en un nuevo T-MAR.

```

1: function registration_serving_tmar(tedmm_entry * tbce)
  ▷ ① Creación de un nuevo túnel entre el T-MAR actual y el previo
2:   tunnel=createNewTunnel(tbce → our_addr, tbce → mn_prev_tmar_addr);
  ▷ ② Añadir una nueva ruta en el T-MAR para el tráfico uplink
3:   addRouteTmar(tbce → mn_prefix, tbce → mn_addr, tunnel);
  ▷ ③ Enviar mensaje RA al MN
4:   sendRAtoMobileNode(tbce → mn_prefix, tbce → mn_link_local_addr);
5: end function

```

- Función “*update_previous_mar*”: esta función entrará en juego cuando un T-MAR reciba un mensaje PBU, es decir, en el T-MAR previo. Permitirá actualizar el registro de un MN que realizó su movimiento y asociación a otro T-MAR. La implementación y funcionamiento se muestra en el Algoritmo 2. Como se ha comentado anteriormente, esta función finalizará la creación del túnel bidireccional, estableciendo un túnel *downlink* hacia el T-MAR actual ①. Posteriormente, se actualizan las rutas en el propio agente de movilidad, para que el tráfico dirigido al MN sea tunelizado desde el T-MAR previo hacia el T-MAR actual ②. Por último, se envía un mensaje PBA al CMD como respuesta al PBU recibido, con toda la información del nuevo vínculo de movilidad ③.

Algoritmo 2 Actualización de un T-MAR que ha visitado el MN previamente.

```

1: function update_previous_mar(tedmm_entry * tbce)
  ▷ ① Creación de un nuevo túnel entre el T-MAR previo y el actual
2:   tunnel=createNewTunnel(tbce → our_addr, tbce → mn_serv_tmar_addr);
  ▷ ② Actualizar la ruta en el T-MAR para el tráfico downlink
3:   updateRoutePrevTmar(tbce → mn_prefix, tbce → mn_addr, tunnel);
  ▷ ③ Enviar mensaje PBA al CMD con el vínculo de movilidad completo
4:   sendPBtoCMD(tbce);
5: end function

```

Una vez que se ha descrito la implementación de TE-DMM y todos los aspectos relacionados con ella, se llevará a cabo la evaluación de rendimiento experimental entre

NB-DMM y la propuesta TE-DMM. Dicha evaluación se ha realizado sobre un *testbed* experimental, cuya configuración se muestra en la Figura 3.15. Ambas soluciones han sido evaluadas sobre la misma topología de red, proporcionando una comparativa aún más fiable.

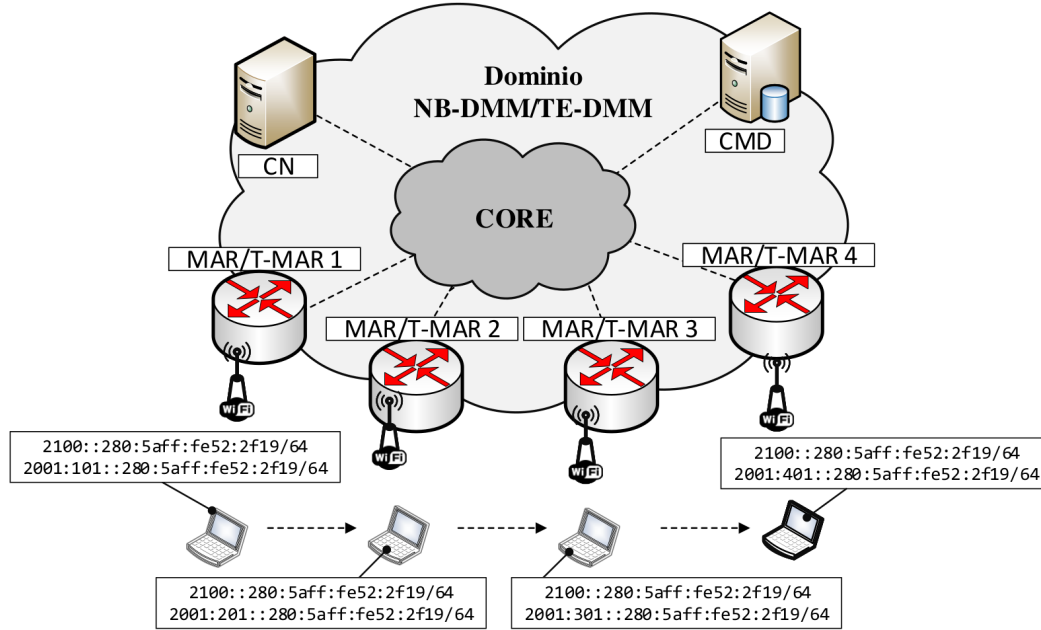


Figura 3.15: *Testbed* usado para la evaluación experimental de TE-DMM y NB-DMM

La red de acceso está formada por cuatro agentes de movilidad (MAR/T-MAR), que proporcionan acceso a los nodos móviles a través de *interfaces* inalámbricas, usando tecnología IEEE 802.11g. Además, el comportamiento de la red troncal es modelado utilizando CORE [113]. Dicha red troncal interconecta todos los agentes de movilidad con el CMD y el CN. Los puntos de acceso (*Cisco Aironet 1130AG Series*) se conectan directamente a los agentes de movilidad y son utilizados por los nodos móviles para conectarse a la red de acceso.

Todos los dispositivos, exceptuando los puntos de acceso, hacen uso de sistemas operativos *open source*. Los MAR, T-MAR y CMD ejecutan un sistema operativo Linux con un *kernel* que ha sido compilado con unas características específicas. Por lo tanto, para que ambas soluciones, NB-DMM y TE-DMM, funcionen correctamente sobre Linux, es necesario compilar el *kernel* añadiendo el soporte a la movilidad IPv6, que no está activado por defecto. Por ello, es necesario iniciar un proceso de configuración del *kernel*, activando las características de movilidad necesaria para, posteriormente,

compilarlo de nuevo y, así, los cambios surtan efecto. Las opciones relacionadas con el soporte de movilidad que tienen que ser habilitadas se muestran a continuación.

Linux Kernel (Características para el soporte a la movilidad)

```

├─ General setup
│   └─ Prompt development and/or incomplete code/drivers
│   └─ System V IPC
├─ Networking support
│   └─ Networking options
│       ├── Transformation user configuration interface
│       ├── Transformation sub policy support
│       ├── Transformation migrate database
│       ├── PF_KEY sockets
│       ├── PF_KEY migrate
│       ├── TCP/IP networking
│       └─ The IPv6 protocol
│           ├── IPv6: AH transformation
│           ├── IPv6: ESP transformation
│           ├── IPv6: IPComp transformation
│           ├── IPv6: Mobility
│           ├── IPv6: IPsec transport mode
│           ├── IPv6: IPsec tunnel mode
│           ├── IPv6: MIPv6 route optimization mode
│           ├── IPv6: IPv6-in-IPv6 tunnel
│           ├── IPv6: Multiple Routing Tables
│           └─ IPv6: source address based routing
├─ File systems
│   └─ Pseudo filesystems
│       └─ /proc file system support

```

Una vez llevado a cabo el proceso descrito anteriormente, se instalan las implementaciones de NB-DMM y TE-DMM en los nodos que actúan como MAR, T-MAR y CMD. Por el contrario, tanto el CN como el MN utilizan un sistema operativo Linux que no requiere ningún cambio en el *kernel* ni *software* adicional.

La evaluación experimental llevada a cabo se ha centrado en la evaluación de la latencia de *handover* y de la pérdida de paquetes en ambas propuestas. La latencia de *handover* ha sido tomada como el intervalo de tiempo en el cual el MN pierde la conectividad IP tras realizar un proceso de *handover*.

En la evaluación experimental se ha usado Wireshark para extraer y analizar los eventos producidos durante la misma. El proceso que se ha seguido ha sido el siguiente: inicialmente, el MN se asocia al MAR1/T-MAR1 y comienza a recibir tráfico UDP desde el CN. Posteriormente, el MN visita MAR2/T-MAR2, MAR3/T-MAR3 y MAR4/T-MAR4 manteniendo activo el flujo inicial asociado al primer agente de

movilidad. Una vez hecho esto, el MN vuelve al MAR1/T-MAR1 y repite el proceso, obteniéndose un total de 300 *handovers* para ambas soluciones. La Figura 3.16 muestra los resultados obtenidos, representando la función de distribución acumulada (CDF) experimental para los valores de tiempo de recuperación de flujos UDP.

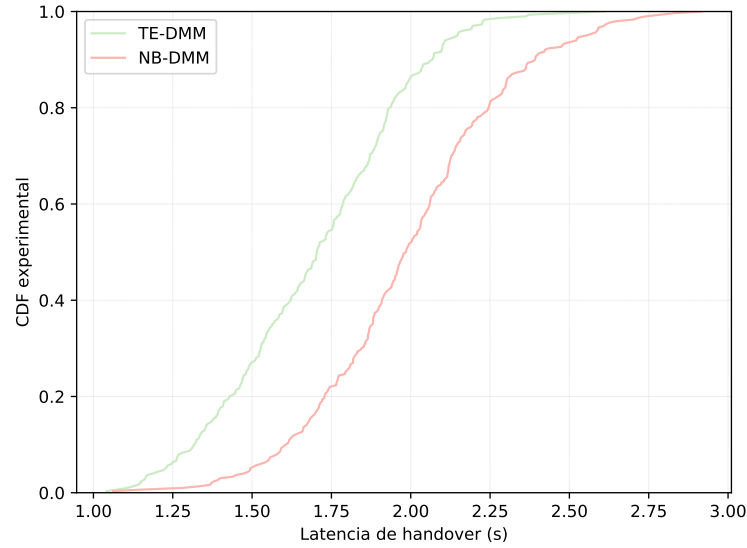


Figura 3.16: Función de distribución acumulada (CDF) experimental de las medidas de latencia *handover*

La Figura 3.17 muestra los paquetes perdidos calculados a través de la Ecuación 3.15, usando la media de todos los valores de latencia de *handover* obtenidos anteriormente para ambas propuestas.

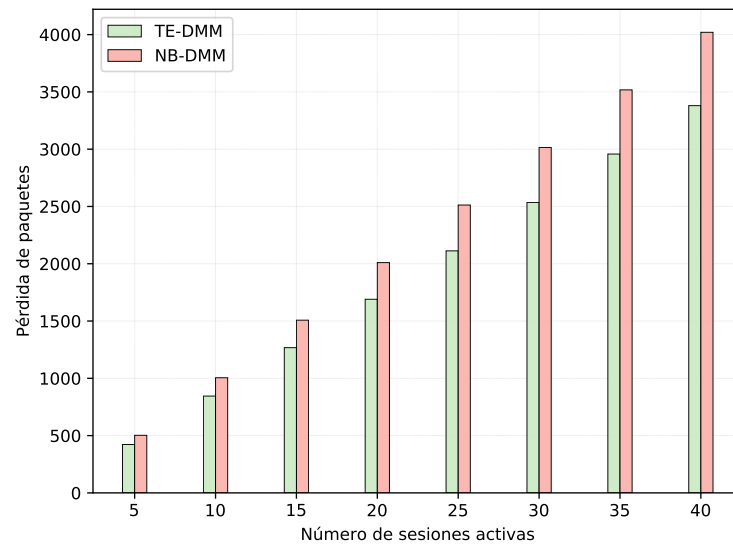


Figura 3.17: Evaluación experimental de paquetes perdidos

Haciendo referencia a la evaluación de los paquetes perdidos (Figura 3.17) en función del número de sesiones activas (ρ_S), TE-DMM presenta un rendimiento superior a NB-DMM, obteniendo unos valores de pérdida de paquetes menores, debido a que no se tiene que realizar la actualización de todas las anclas de movilidad anteriores, simplemente se realiza una extensión del túnel existente. Esto se traduce en una mejora de latencia de *handover* y, por consiguiente, en una reducción de la pérdida de paquetes en el MN.

3.5. Conclusiones

En este capítulo se presenta un enfoque novedoso de gestión de la movilidad distribuida, llamado TE-DMM, que permite una gestión eficiente del plano de control, obteniendo mejoras significativas del tráfico de señalización, entre otras.

Se ha llevado a cabo una evaluación de rendimiento analítica y experimental entre la propuesta TE-DMM y la solución más representativa de gestión de la movilidad distribuida (NB-DMM). El modelo analítico ha permitido comparar ambas propuestas, poniendo de manifiesto los beneficios introducidos por TE-DMM en términos de costes de señalización, los cuales son reducidos aproximadamente en un 50 %, con respecto a NB-DMM. Además, el análisis revela que TE-DMM reduce la complejidad del plano de control, mejorando el proceso de gestión de túneles durante el movimiento de los usuarios.

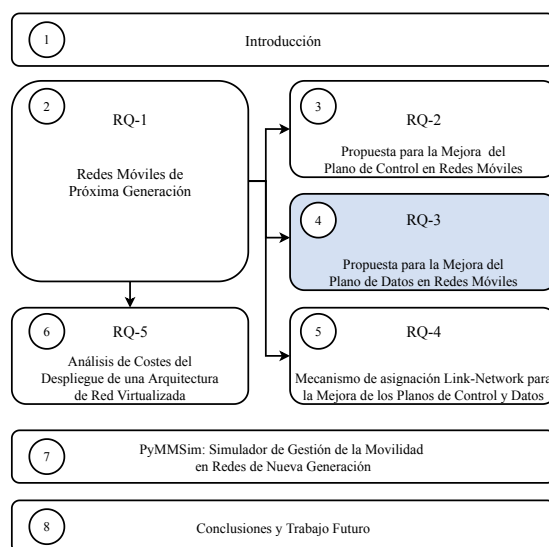
Por otro lado, para realizar la evaluación experimental se ha llevado a cabo la implementación de nuestra propuesta en un sistema Linux. Por lo tanto, se han utilizado implementaciones reales de NB-DMM y nuestro mecanismo (TE-DMM) sobre un *testbed* de movilidad real. Los resultados obtenidos muestran que TE-DMM mejora la latencia de *handover*, proporcionando, además, beneficios significativos en relación a la pérdida de paquetes durante dicho proceso.

Se puede concluir, por tanto, que TE-DMM realiza un mejor uso de los recursos de la red y mejora el comportamiento del plano de control con respecto a NB-DMM.

Capítulo 4

Propuesta para la Mejora del Plano de Datos en Redes Móviles

El tráfico de datos móviles está creciendo exponencialmente [18], debido al incremento de los dispositivos inteligentes inalámbricos y la cantidad de servicios que las redes actuales pueden admitir. Este crecimiento significativo está impulsando a los operadores de redes móviles a proponer nuevas soluciones para mejorar el rendimiento y la eficiencia de su red. Por ello, en este capítulo se proporciona una solución de gestión de la movilidad distribuida, llamada *Software Redirection to Distributed Mobility Management* (SR-DMM), que hace uso de las capacidades y beneficios que aporta SDN, con el objetivo de mejorar el rendimiento del plano de datos, proporcionando flexibilidad, escalabilidad y fiabilidad a las redes móviles de nueva generación.



El rendimiento del mecanismo propuesto SR-DMM ha sido evaluado de manera analítica y experimental, con respecto a la solución más representativa de gestión de la movilidad distribuida basada en la red (NB-DMM), cuya operación se describe de manera detallada en la Sección 2.3.2, y la visión general de la arquitectura se muestra en la Figura 4.1. En la evaluación de rendimiento se presta especial atención al coste de señalización y al coste de entrega de paquetes; además de llevar a cabo un análisis y evaluación experimental de la latencia de *handover*.

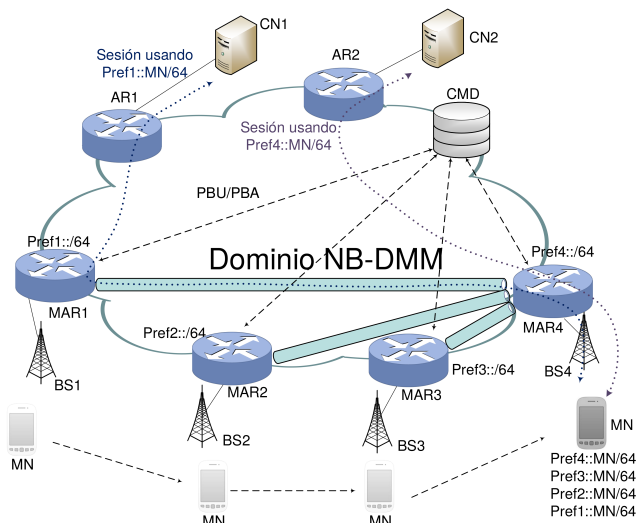


Figura 4.1: Arquitectura de NB-DMM

En definitiva, el capítulo se organiza tal y como se presenta a continuación:

- La Sección 4.1 describe de manera exhaustiva el mecanismo propuesto para la gestión eficiente del plano de datos en redes móviles, llamado SR-DMM.
- En la Sección 4.2 se desarrolla un modelo analítico que permite evaluar el rendimiento del mecanismo SR-DMM con respecto a la solución más representativa de gestión de la movilidad distribuida (NB-DMM). Además, se definen las métricas utilizadas para la evaluación de la propuesta.
- En la Sección 4.3 se presentan y comparan analíticamente los resultados obtenidos para ambas soluciones (NB-DMM y SR-DMM).
- La Sección 4.4 describe la implementación desarrollada del mecanismo SR-DMM y presenta una evaluación experimental comparativa de ambas soluciones sobre un *testbed* de movilidad, usando implementaciones reales en ambos casos, con el objetivo de realizar una validación experimental.

- La Sección 4.5 muestra las conclusiones que han sido obtenidas a partir del análisis realizado y la evaluación llevada a cabo.

4.1. Enfoque de gestión de movilidad basado en SDN

Impulsados por el aumento masivo del tráfico de datos móviles [18], los mecanismos eficientes de gestión y administración de redes se han revelado como uno de los principales desafíos en las redes móviles de próxima generación [114, 115]. Específicamente, en estos últimos años, la tecnología SDN está siendo considerada como una oportunidad importante para las arquitecturas de redes móviles 5G [42]. Específicamente en este contexto y, tal y como se explicó en el Capítulo 2, SDN ha tomado gran importancia en el diseño de soluciones de gestión de la movilidad IP [80, 96, 98], con el objetivo de poder controlar y asignar dinámicamente los recursos de red para proporcionar flexibilidad, escalabilidad y fiabilidad en este nuevo ecosistema 5G [4].

Tal y como se muestra en la Figura 4.2, la arquitectura funcional de SDN está formada por tres niveles principalmente: el plano de datos, que contiene el equipamiento físico de red; el plano de control, constituido por los controladores de red y el plano de aplicación, donde se despliegan las aplicaciones funcionales de red. Además, en esta Figura 4.2 también se muestra cómo se desplegaría una aplicación de gestión de la movilidad sobre la propia arquitectura funcional de SDN, que constituye el objetivo principal del presente capítulo.

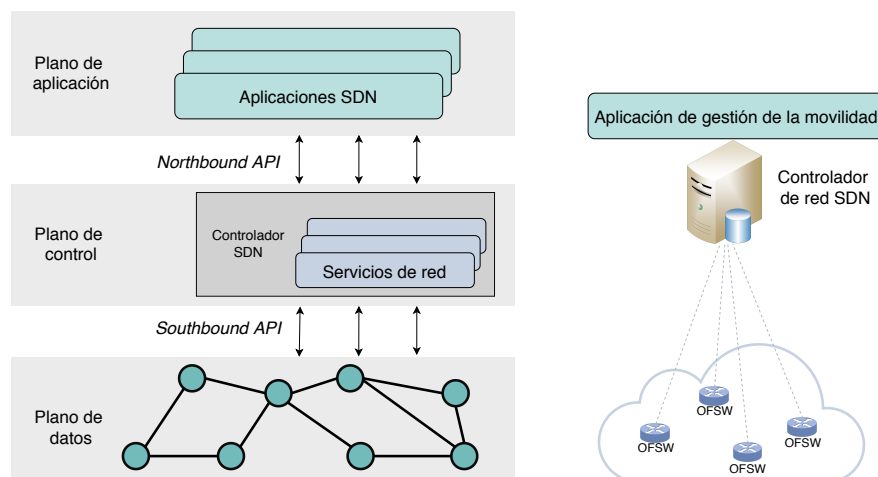


Figura 4.2: Arquitectura funcional de SDN

En definitiva, SDN proporciona innovación, un rendimiento superior y configuraciones de red mejoradas con respecto a la gestión en redes tradicionales. Por

ello, esta tecnología es considerada como una gran oportunidad para la gestión de la movilidad eficiente en redes móviles.

A continuación, se describe de manera detallada la operación y funcionamiento del segundo mecanismo propuesto en esta Tesis, cuyo principal objetivo es llevar a cabo una gestión eficiente del plano de datos (SR-DMM).

4.1.1. Solución SR-DMM

El propósito de gestionar la movilidad de los usuarios móviles, haciendo uso de la tecnología SDN, es el de conseguir soluciones donde los planos de datos y de control estén separados y distribuidos por la red. En este contexto, el soporte a la movilidad IP se proporciona como un servicio desarrollado en *software*, en forma de aplicación SDN, que se ejecuta en el controlador de red, y administra un plano de control formado por *hardware* de red genérico. De este modo, el principal objetivo de nuestra propuesta SR-DMM es proporcionar flexibilidad, escalabilidad y fiabilidad a las redes móviles de nueva generación, aprovechando las capacidades de la arquitectura funcional de la tecnología SDN, ofreciendo el soporte a la movilidad como un servicio y aportando eficiencia a nivel de flujos de datos IPv6. La flexibilidad de red se refiere a la capacidad de la red para adaptarse y gestionar sus recursos convenientemente [71, 116]. Esta flexibilidad de red es ofrecida por la propuesta SR-DMM gracias a la programabilidad de SDN. La solución propuesta implementa un mecanismo a través de estándares abiertos para redireccionar flujos eficientemente cuando un nodo móvil se mueve dentro del dominio de red. Por otro lado, la escalabilidad es definida como la capacidad para manejar cargas de trabajo cada vez mayores, más específicamente en el plano de control [72]. Dicha escalabilidad es proporcionada por SR-DMM a través de la reducción de la sobrecarga de señalización, ya que no es requerido el proceso de actualización de vínculos de movilidad. En las siguientes secciones, la flexibilidad y escalabilidad son evaluadas analítica y experimentalmente en términos de costes de señalización, costes de entrega de paquetes y latencia de *handover*.

Las arquitecturas SDN cuentan con un controlador centralizado que tiene una visión global de la red, de la cual se beneficiará la aplicación de gestión de la movilidad desarrollada (SR-DMM), a través del protocolo OpenFlow [33]. De esta forma, el servicio de soporte a la movilidad podrá conocer el estado global del plano de datos, pudiendo tomar decisiones en tiempo real y proporcionando además la flexibilidad requerida en entornos de movilidad 5G. Además, SR-DMM dota al controlador con la capacidad para anclar los flujos IPv6 de los usuarios móviles a los *switches* del borde de la red de acceso, a través de la interfaz OpenFlow. Por tanto, la solución

propuesta distingue entre *switches* con capacidad de anclar flujos IPv6, que son los *switches* ubicados en el borde de la red de acceso, y *switches* sin capacidad para anclar flujos de red. La aplicación SR-DMM es desarrollada en el controlador de red y permite configurar tablas de flujo sobre *switches* OpenFlow (OFSwitches), proporcionándoles además capacidades de descubrimiento de vecinos [117] y control de acceso.

A diferencia de lo que ocurre en soluciones NB-DMM, la propuesta SR-DMM evita sobrecargas introducidas por el mecanismo de tunelización entre las anclas de movilidad, debido a la visión general del controlador de red, que mantiene actualizada en todo momento la localización de cada MN. Por lo tanto, SR-DMM utiliza redirección de flujos de paquetes IPv6, en lugar de mecanismos de tunelización IP sobre IP.

La Figura 4.3 muestra la arquitectura funcional de la propuesta SR-DMM. Tal y como se puede observar, la red de acceso está formada por *switches* OpenFlow, que son gestionados por el controlador de red SDN. En este caso concreto, un *switch* ubicado en el borde de la red de acceso (OFSwitch1) actúa como ancla para un flujo de paquetes IPv6 iniciado cuando el MN se encontraba conectado a dicho *switch*. Cuando el MN realiza su movimiento por el dominio SR-DMM, adquiere nuevos prefijos de red IPv6 en las redes visitadas.

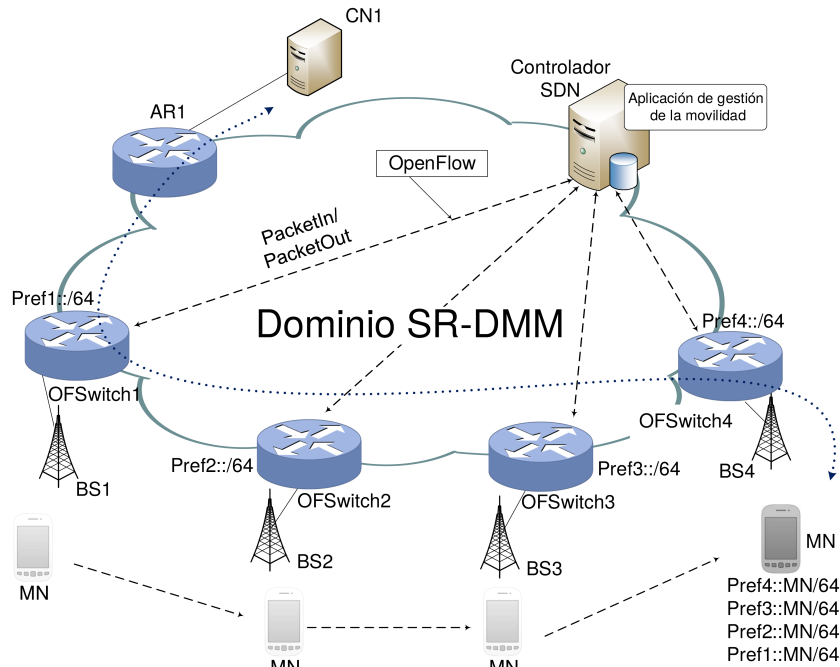


Figura 4.3: Arquitectura funcional de la propuesta SR-DMM

Sin duda, se puede establecer una analogía entre la propuesta SR-DMM y la solución NB-DMM; el controlador SDN en SR-DMM es una entidad similar al CMD (NB-DMM)

y los *switches* OpenFlow (OFSwitches) son las entidades análogas a los MAR de NB-DMM ubicados en el borde de la red de acceso.

En las siguientes subsecciones se describe la operación de las tareas más importantes llevadas a cabo por la propuesta SR-DMM: el registro inicial de un nodo móvil en el dominio de red y el proceso de *handover*.

4.1.2. Registro de un MN con SR-DMM

En la propuesta SR-DMM, una vez que el MN se conecta por primera vez a un OFSwitch del dominio, envía un mensaje *Router Solicitation* (RS) al OFSwitch de acceso ①. Este OFSwitch, según su tabla de flujos, encapsula el RS en un mensaje *OpenFlow PacketIn*¹ y lo envía al controlador de red SDN ②. Cuando llega al controlador, este se encarga de autenticar y localizar al MN, creando y almacenando el nuevo vínculo de movilidad con su dirección MAC, el identificador del propio nodo móvil y el ancla de movilidad a la que está actualmente asociado el terminal. A continuación, el controlador SDN construye un mensaje *Router Advertisement* (RA) con el nuevo prefijo de red, lo encapsula en un mensaje *OpenFlow PacketOut* y finalmente lo envía al OFSwitch ③. Cuando este lo recibe, desencapsula el mensaje RA y lo envía al MN ④, que cuando lo recibe autoconfigura su dirección IPv6, quedando anclado al OFSwitch con el prefijo de red correspondiente. La Figura 4.4 muestra el diagrama de flujo de mensajes del registro inicial de un MN en el dominio SR-DMM.

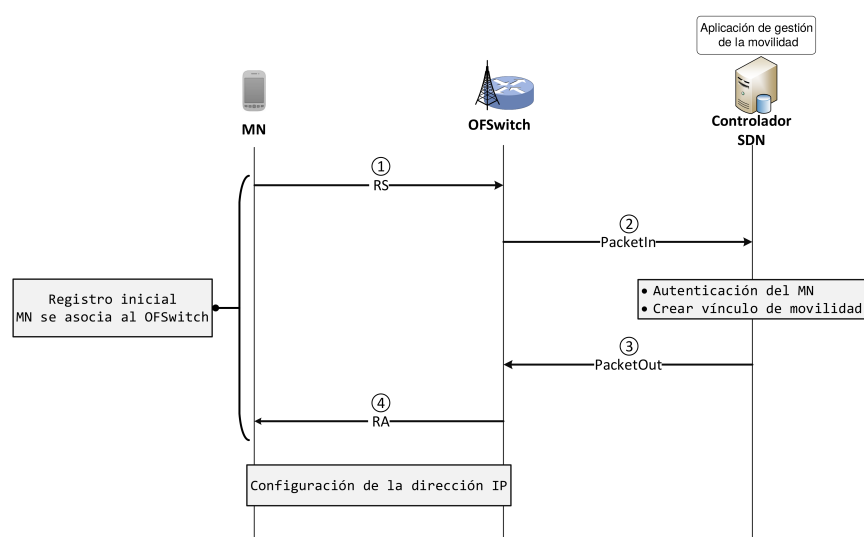


Figura 4.4: Diagrama de flujo de mensajes: registro inicial en SR-DMM

¹Más información de los mensajes del protocolo OpenFlow: <http://flowgrammable.org/sdn/openflow/>

4.1.3. Proceso de *handover* con SR-DMM

Con respecto a la operación durante el proceso de *handover*, la aplicación SR-DMM permite detectar la ubicación del MN cuando el controlador SDN recibe un mensaje RS ① encapsulado dentro de un mensaje *PacketIn* ②. Cuando el controlador SDN recibe el *PacketIn* verifica que el OFSwitch actual es distinto del OFSwitch previo, obteniendo tantos prefijos de red como OFSwitches previos haya visitado el MN durante su movimiento por el dominio. A través de estos prefijos de red, el nodo móvil puede autoconfigurar sus direcciones IPv6, una vez haya recibido el mensaje RA desde el nuevo OFSwitch que actúa como agente de movilidad. La Figura 4.5 muestra el diagrama de flujo de mensajes del proceso de *handover* de un MN en el dominio SR-DMM.

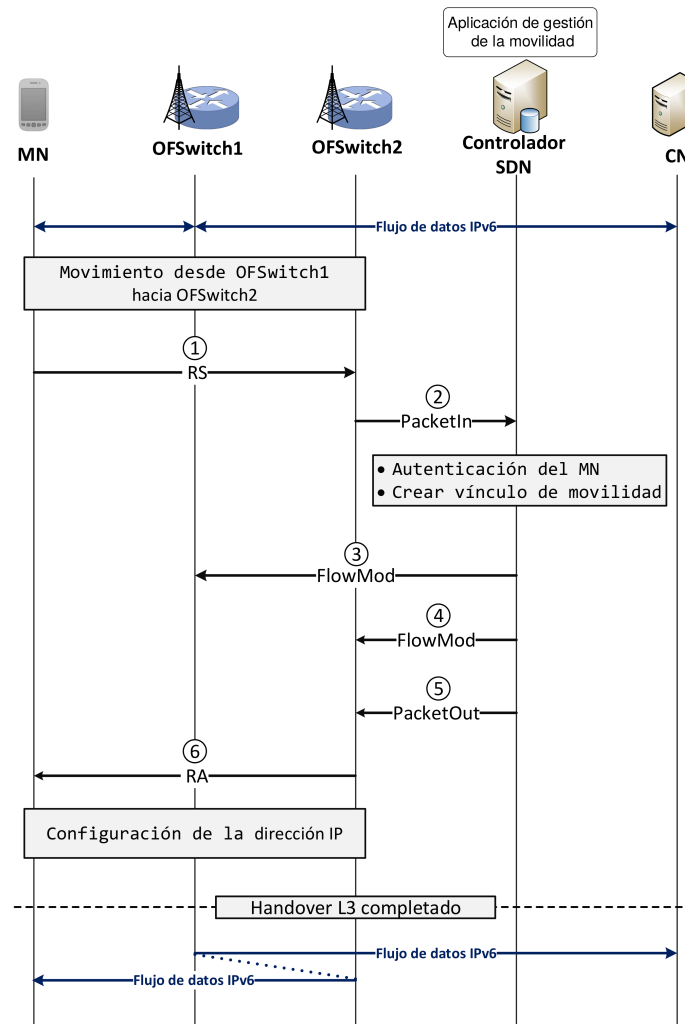


Figura 4.5: Diagrama de flujo de mensajes: proceso de *handover* en SR-DMM

La aplicación SR-DMM instala una regla de flujo en cada OFSwitch previo donde están anclados los flujos iniciados por el MN, a través de mensajes *OpenFlow FlowMod* enviados a cada uno de ellos ③. La creación de las reglas de flujos para los OFSwitches previos visitados por un nodo móvil es descrita en detalle en el Algoritmo 3.

Algoritmo 3 Creación de las reglas de flujo para los OFSwitches visitados por el MN.

Entrada:

Nodo móvil MN ,
 Lista de OFSwitches previos L_{OFS} ,
 OFSwitch actual A_{OFS}

Salida:

Diccionario de reglas de flujo a instalar en los OFSwitches previos L_{FR}

```

1: for  $P_{OFS} \in L_{OFS}$  do
  ▷ Construir la dirección IPv6 previa del MN
2:    $MN_{prev\_addr} = \text{getIPv6Address}(MN \rightarrow hw\_addr, P_{OFS} \rightarrow prefix);$ 
  ▷ Obtener la dirección IPv6 que se utilizará para establecer conexión entre el
  OFSwitch previo y el actual
3:    $MN_{local\_addr} = \text{getIPv6Address}(A_{OFS} \rightarrow prefix, MN \rightarrow id);$ 
  ▷ Construir regla OpenFlow para instalar en los OFSwitches previos
4:    $R = \text{buildOpenFlowRule}(P_{OFS}, MN_{prev\_addr}, MN_{local\_addr});$ 
  ▷ Insertar la regla creada en el diccionario que contiene todas las reglas OpenFlow
  creadas por el algoritmo
5:    $L_{FR}\{P_{OFS}\} = R;$ 
6: end for
7: return  $L_{FR};$ 

```

Las reglas instaladas en cada OFSwitch previo permite redirigir los paquetes desde las anclas de movilidad visitadas con anterioridad hacia el OFSwitch actual, al que se encuentra conectado el MN; estableciendo como dirección IPv6 destino la formada por el prefijo de red asociado al OFSwitch actual y el identificador del MN en el dominio SR-DMM.

Por otro lado, es necesario destacar que también es importante instalar una nueva regla de flujo en el agente de movilidad actual. Se trata de una operación simple, que tiene como único objetivo restaurar la dirección IPv6 original de los paquetes pertenecientes a las sesiones anteriores, para mantener la continuidad de las mismas de forma transparente. De este modo, el MN recibirá todo el flujo de datos convenientemente, como si estuviera conectado al ancla inicial. La creación de esta regla es llevada a cabo en el controlador, el cual la encapsula en un mensaje *FlowMod* y se la envía al OFSwitch actual ④, tal y como se muestra en la Figura 4.5. Por último, tal y como ocurría en el proceso de registro inicial (ver subsección 4.1.2), el controlador

SDN construye un mensaje *Router Advertisement* (RA) con el nuevo prefijo de red, lo encapsula en un mensaje *OpenFlow PacketOut* y lo envía al OFSwitch ⑤. Cuando este lo recibe, desencapsula el mensaje RA y lo envía al MN ⑥, que cuando lo recibe autoconfigura su dirección IPv6, quedando anclado al OFSwitch con el prefijo de red correspondiente.

Además, en SR-DMM las reglas de flujo instaladas en OFSwitches expiran en un periodo de tiempo determinado si no existen las coincidencias correspondientes. El controlador de red es el encargado de gestionar esta operación, de tal manera que las sesiones que hayan expirado quedarán sin soporte de movilidad.

En definitiva, la propuesta SR-DMM permite que el tráfico de datos se redirija del OFSwitch previo al OFSwitch actual de forma transparente al usuario, sin necesidad de involucrar al propio MN en cuestiones relativas al plano de control (señalización) y utilizando la ubicación del MN para tal efecto. De esta forma, SR-DMM no introduce sobrecarga adicional de señalización en la red, ya que evita el uso de túneles IP sobre IP, a diferencia de otras soluciones como NB-DMM. Existen otras implementaciones en la literatura que establecen una ruta óptima mediante la aplicación de reglas de flujo a todos los OFSwitches involucrados en la propia ruta entre el CN y el MN [96]. Sin embargo, con SR-DMM solamente es necesario instalar nuevas reglas en los OFSwitches del borde de la red de acceso visitados por el MN. Esto es así debido a que las rutas hacia cada subred de destino se calculan de manera dinámica por el servicio de movilidad, obteniendo la ruta óptima entre el OFSwitch previo y el actual en cada momento. Además, todos los dispositivos de red del dominio SR-DMM son genéricos y su funcionalidad es establecida por el controlador SDN, de tal manera que no es necesario utilizar ningún agente de movilidad con capacidades especiales.

4.2. Modelo analítico

En esta sección se desarrolla un modelo analítico para evaluar el rendimiento de la solución de gestión de la movilidad distribuida (NB-DMM) y nuestra propuesta basada en SDN (SR-DMM) de forma comparativa, en términos de coste de señalización y coste de entrega de paquetes.

El análisis de rendimiento se realiza en un dominio de red compuesto con N celdas de cobertura conectadas a diferentes nodos de acceso (OFSwitches de borde), que son los primeros nodos con capacidad IP del dominio. La Figura 4.6 muestra las topologías de red de ambas soluciones (NB-DMM y SR-DMM) utilizadas en la evaluación analítica.

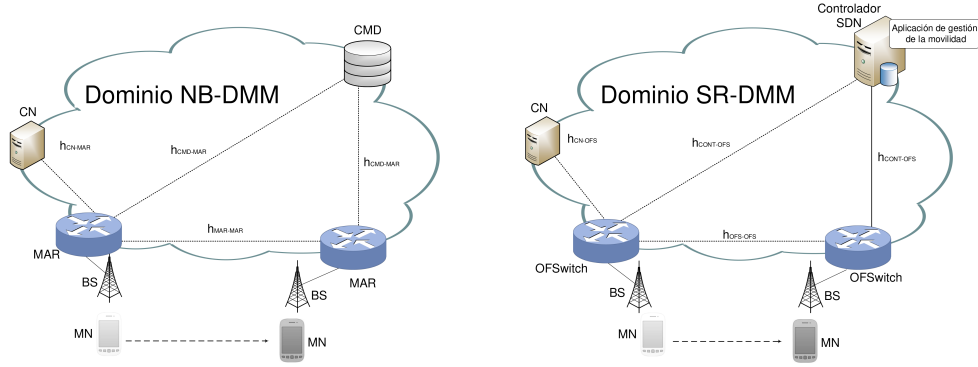


Figura 4.6: Topología utilizada para el modelo analítico de NB-DMM y SR-DMM

El modelo analítico ha sido desarrollado usando el *framework* descrito en [107, 85]. El coste de transmisión de paquetes en redes IP es directamente proporcional al número de saltos entre los nodos origen y destino. Por ello, en la Figura 4.6 se han definido las distancias entre las entidades más importantes de cada solución, que luego serán utilizadas durante el desarrollo del modelo analítico, estableciéndose h_{x-y} como la distancia en número de saltos entre los nodos de red x e y .

4.2.1. Coste de señalización

Una de las principales funcionalidades para cualquier protocolo de gestión de movilidad IP es asegurar que la sesión de movilidad de cada MN se mantenga activa y actualizada mientras él realiza su movimiento por el dominio de red. Esto requiere de mensajes de control que deben enviarse entre los agentes de movilidad en la red.

Analizando el plano de control de las propuestas estudiadas en este capítulo, se puede definir el coste total de señalización relativo a la actualización de vínculos de movilidad durante una sesión como C_S , el cual representa la carga de tráfico acumulada en el intercambio de señalización. Este coste depende del tamaño de los propios mensajes de señalización y del número de *handovers* de nivel 3 realizados durante el intervalo de tiempo en el que la comunicación en el MN se mantiene activa. Para cada movimiento en la solución NB-DMM, se envía un mensaje *Proxy Binding Update* (PBU) y uno de respuesta *Proxy Binding Acknowledgement* (PBA). Además, en NB-DMM, hay que tener en cuenta la actualización de la tabla de vínculos de movilidad durante el tiempo de vida del prefijo de red, así como la eliminación del prefijo una vez que no exista ninguna sesión activa.

Por otro lado, la propuesta SR-DMM notifica el *handover* usando mensajes OpenFlow, interactuando con las tablas de flujos de los distintos OFSwitches. Los

mensajes de señalización intercambiados son mensajes *FlowMod*, típicos en cualquier arquitectura SDN. De esta manera, la ruta óptima entre agentes de movilidad es actualizada a través de mensajes *FlowMod* de OpenFlow. Además, los prefijos de red inactivos son eliminados al expirar los temporizadores de las entradas de la tabla de flujo de movilidad de los OFSwitches, no siendo necesarios los mecanismos de refresco utilizados por NB-DMM.

En definitiva, definimos el coste total de señalización como la suma de tres componentes principales: el coste de actualización del vínculo de movilidad después de un *handover*, el coste de eliminación de un prefijo de red que ha expirado y el coste requerido para actualizar periódicamente los vínculos de movilidad. Las siguientes expresiones (Ecuaciones 4.1 y 4.2) muestran el coste de señalización durante el movimiento del MN para la solución NB-DMM ($C_S(NB - DMM)$) y la propuesta SR-DMM ($C_S(SR - DMM)$).

$$\begin{aligned} C_S(NB - DMM) = & \mu_c \cdot ((S_{PBU} + S_{PBA}) \cdot h_{CMD-MAR} \\ & \cdot (N_{pr} + 1) + 2 \cdot (S_{PBU} + S_{PBA}) \cdot h_{CMD-MAR} \\ & + R_{BCE} \cdot (S_{PBU} + S_{PBA}) \cdot h_{CMD-MAR}) \end{aligned} \quad (4.1)$$

$$\begin{aligned} C_S(SR - DMM) = & \mu_c \cdot ((N_{pr} + 1) \cdot 2 \cdot S_{FMOD} \cdot h_{CONT-OFS} \\ & + (S_{PIN} + S_{RS} + S_{POUT} + S_{RA}) \cdot h_{CONT-OFS}) \end{aligned} \quad (4.2)$$

donde μ_c es la tasa de traspaso de una región de cobertura a otra y N_{pr} representa el número de prefijos activos por MN. N_{pr} puede ser definido como el número de MAR/OFSwitch que mantienen alguna sesión activa con el MN. Tal y como se define en [88], N_{pr} es calculado como se muestra en la Ecuación 4.3.

$$N_{pr} = \frac{\mu_c}{\delta} \quad (4.3)$$

donde $1/\delta$ representa el tiempo de vida de un prefijo de red mientras el MN visita una red determinada.

Por otro lado, se asume que el tiempo medio de estancia en una celda determinada T_c para un MN es una variable aleatoria que sigue una distribución exponencial [78], y puede calcularse tal y como se muestra en la Ecuación 4.4.

$$T_c = \frac{1}{\mu_c} \quad (4.4)$$

La Tabla 4.1 recoge todos los parámetros utilizados para el modelo analítico y el análisis de rendimiento, así como una breve descripción de cada uno de ellos.

Tabla 4.1: Notación utilizada en el modelo analítico

Parámetros
h_{x-y} : distancia en número de saltos entre los nodos x e y .
N_{pr} : número de prefijos activos usados.
$N_{p/s}$: tasa de transmisión de paquetes por flujo activo.
R_{BCE} : tasa de refresco del vínculo de movilidad.
S_{DATA} : tamaño medio de los paquetes de datos.
S_{PBU} : tamaño medio de un mensaje PBU.
S_{PBA} : tamaño medio de un mensaje PBA.
S_{FMOD} : tamaño medio del mensaje <i>FlowMod</i> .
S_{PIN} : tamaño medio del mensaje <i>PacketIn</i> .
S_{POUT} : tamaño medio del mensaje <i>PacketOut</i> .
S_{RS} : tamaño del mensaje <i>Router Solicitation</i> .
S_{RA} : tamaño del mensaje <i>Router Advertisement</i> .
S_{IP} : tamaño de la cabecera de tunelización IPv6.
μ_c : tasa de traspaso entre celdas.
T_c : tiempo medio de estancia en una celda determinada.
δ : tiempo de vida de un prefijo de red.

4.2.2. Coste de entrega de los paquetes de datos

El coste total de entrega de paquetes de datos para una sesión es definido como C_{PD} . Este valor está influenciado por el tamaño medio de los mensajes de datos multiplicado por el número de saltos necesarios para reenviar los paquetes desde el CN hasta el MN, y viceversa. Por lo tanto, las expresiones que representan el coste C_{PD} para la solución NB-DMM y la propuesta SR-DMM, se muestran en las Ecuaciones 4.5 y 4.6, respectivamente.

$$\begin{aligned}
 C_{PD}(NB - DMM) = N_{p/s} \cdot ((N_{pr} - 1) \cdot (S_{DATA} + S_{IP}) \\
 \cdot h_{MAR-MAR} + (S_{DATA} \cdot h_{CN-MAR}) \\
 + (S_{DATA} \cdot h_{MN-MAR}))
 \end{aligned} \quad (4.5)$$

$$\begin{aligned}
C_{PD}(SR - DMM) = N_{p/s} \cdot ((N_{pr} - 1) \cdot S_{DATA} \cdot h_{OFS-OFS} \\
\cdot h_{OFS-OFS} + (S_{DATA} \cdot h_{CN-OFS}) \\
+ (S_{DATA} \cdot h_{MN-OFS}))
\end{aligned} \tag{4.6}$$

donde $N_{p/s}$ representa la tasa de transmisión de paquetes por flujo activo y S_{DATA} es el tamaño medio de estos paquetes de datos. Todos los parámetros usados en el análisis y modelo analítico se muestran en la Tabla 4.1.

4.3. Resultados obtenidos

En esta sección se presentan y discuten los resultados numéricos obtenidos de manera analítica, llevando a cabo una primera evaluación de rendimiento entre la solución NB-DMM y nuestra propuesta SR-DMM, en términos de coste de señalización y coste de entrega de paquetes. De esta manera se ha estudiado el impacto de diferentes parámetros sobre los costes de movilidad más representativos. La topología de red usada para esta evaluación analítica se muestra en la Figura 4.6.

Además, a continuación se definen los parámetros por defecto que constituyen las configuraciones básicas de la topología, del modelo de movilidad y del modelo de tráfico usadas en el análisis [76, 85, 118].

- Tiempo medio de estancia en una celda: $T_c = [50 - 1800]$ s.
- Tiempo medio de vida de los prefijos de red activos: $E[\delta] = 60$ s.
- Tiempo medio de refresco de la tabla de vínculos de movilidad: $E[R_{BCE}] = 60$ s.
- Tasa de transmisión de paquetes: $N_{p/s} = 3000$.
- Tamaño de los mensajes de control: $S_{PBU} = S_{PBA} = 76$ bytes.
- Tamaño de los mensajes RS y RA: $S_{RS} = S_{RA} = 52$ bytes.
- Tamaño de los mensajes *PacketIn*, *PacketOut* y *FlowMod* de OpenFlow: $S_{PIN} = 92$ bytes, $S_{POUT} = 103$ bytes y $S_{FMOD} = 116$ bytes.
- Tamaño de los paquetes de datos: $S_{DATA} = 120$ bytes.
- Tamaño de la cabecera de tunelización IPv6: $S_{IP} = 40$ bytes.

- Número de saltos entre las entidades de red: $h_{CMD-MAR} = h_{CONT-OFS} = 2$ saltos, $h_{MAR-MAR} = h_{OFS-OFS} = 2$ saltos, $h_{CN-MAR} = h_{CN-OFS} = 5$ saltos, $h_{MN-MAR} = h_{MN-OFS} = 1$ salto.

La Figura 4.7 muestra la comparativa del coste de señalización de ambas soluciones dependiendo del tiempo de estancia de un MN en una celda determinada, que varía tomando valores entre 50 y 1800 segundos. Tal y como puede observarse, ambas propuestas presentan un comportamiento similar, reportando valores mayores de C_S cuando el tiempo de estancia en la celda es menor.

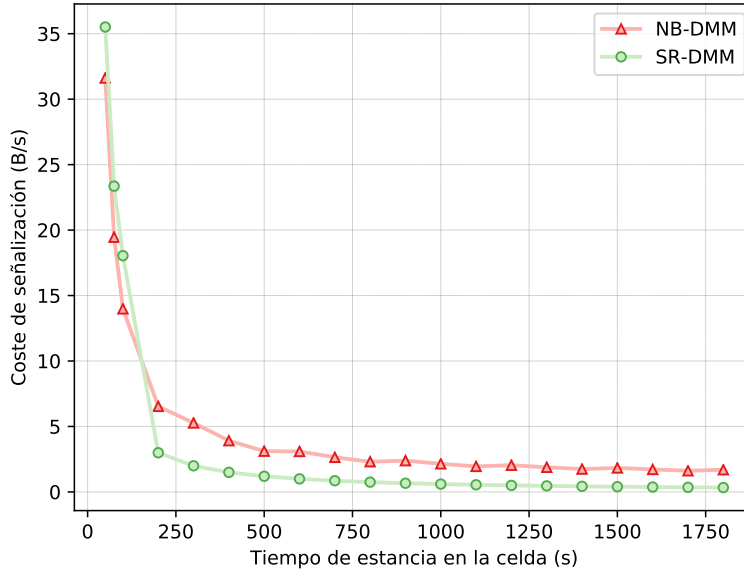


Figura 4.7: Coste de señalización con respecto al tiempo de estancia en una celda

Sin embargo, puede apreciarse que cuando el tiempo de estancia en la celda es bajo, el coste de señalización de la propuesta SR-DMM es superior a los valores obtenidos para la solución NB-DMM. Esto es debido a que valores pequeños de tiempo de estancia en la celda no permiten obtener los beneficios que proporciona SR-DMM, al evitar el proceso de actualización de vínculos de movilidad. Además, SR-DMM es una propuesta basada en SDN y el intercambio de mensajes se lleva a cabo a través de primitivas OpenFlow, que usan el protocolo TCP. Cuando se incrementa el tiempo de estancia en la celda, se observa un rendimiento superior de la propuesta SR-DMM, apreciándose los beneficios aportados al no requerir el proceso de refresco de los vínculos de movilidad.

Por otro lado, el coste de entrega de paquetes representa el coste que supone entregar los paquetes al nodo móvil por unidad de tiempo. La Figura 4.8 muestra el coste de

entrega de paquetes en función del tiempo de estancia en una celda determinada. Como puede observarse, la solución NB-DMM implica tunelización IP sobre IP entre el agente de movilidad actual y los agentes de movilidad previos, incluyendo una nueva cabecera de encapsulación IPv6. A diferencia de NB-DMM, SR-DMM no introduce cabeceras IPv6 adicionales, suponiendo una ventaja importante en el plano de datos. Tal y como se puede observar en la Figura 4.8, la propuesta SR-DMM obtiene mejores resultados de coste de entrega de paquetes, proporcionando una solución de gestión de la movilidad distribuida basada en SDN con el plano de datos optimizado.

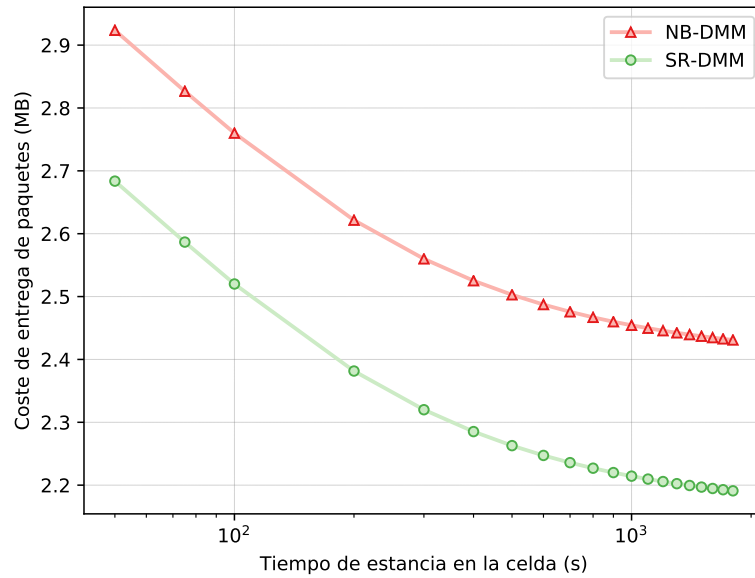


Figura 4.8: Coste de entrega de paquetes con respecto al tiempo de estancia en una celda

4.4. Evaluación experimental

Esta sección presenta una evaluación experimental comparativa utilizando implementaciones reales de NB-DMM y nuestra propuesta SR-DMM.

En el caso de NB-DMM se ha utilizado una implementación para sistemas *GNU/Linux* conocida como MAD-PMIPv6 (*Mobility Anchors Distribution for Proxy Mobile IPv6*) [111, 112], la cual está desarrollada en código C sobre el *kernel* de Linux.

Por otro lado, el plano de control de la propuesta SR-DMM es implementado a través del *framework* de SDN Ryu [119], empleando como interfaz de control el protocolo OpenFlow [33]. Ryu es totalmente compatible con IPv6. Además, el servicio SR-DMM

es desplegado como una aplicación en Python que se ejecuta en el controlador de red, sobre el que está corriendo Ryu.

El *testbed* desplegado para llevar a cabo la evaluación experimental de ambas soluciones se muestra en la Figura 4.9. Ambas soluciones han sido evaluadas sobre la misma topología de red, proporcionando una comparativa aún más fiable.

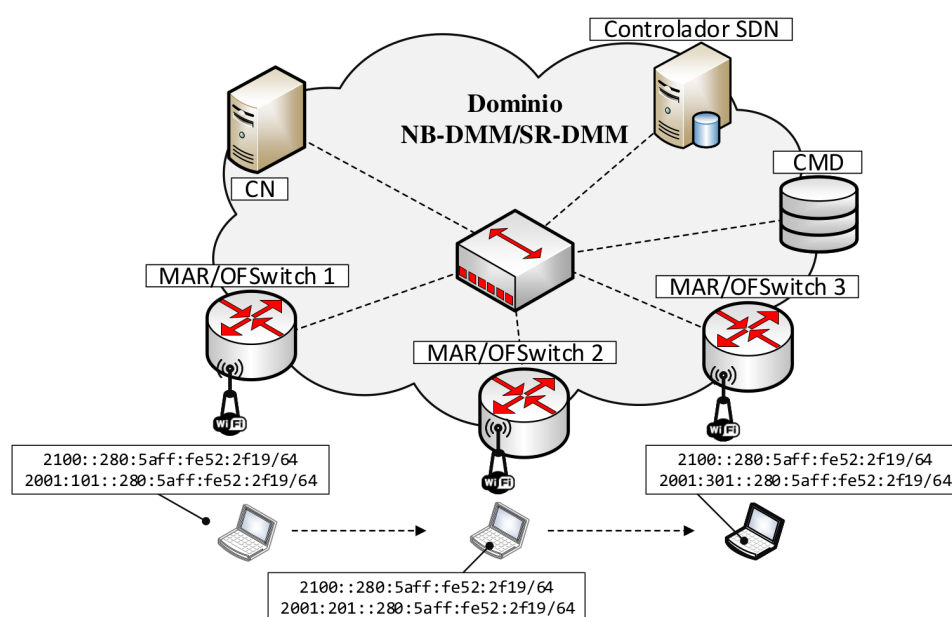


Figura 4.9: Topología del *testbed* utilizado en la evaluación experimental

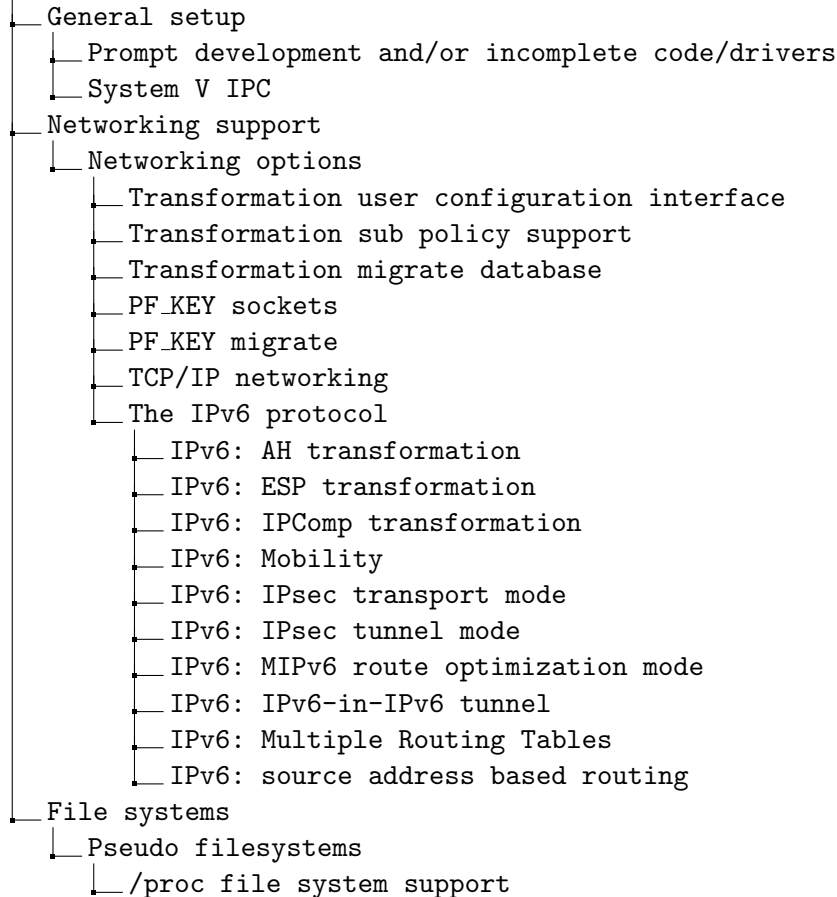
La red de acceso está formada por tres agentes de movilidad (MAR/OFSwitch), que proporcionan acceso a los nodos móviles a través de *interfaces* inalámbricas, usando tecnología IEEE 802.11g. Dicha red troncal interconecta todos los agentes de movilidad con el CMD, el controlador de SDN y el CN. Este último es el que se encarga de enviar flujos IPv6 al MN. Además, los puntos de acceso (*Cisco Aironet 1130AG Series*) se conectan directamente a los agentes de movilidad y son utilizados por los nodos móviles para conectarse a dicha red de acceso.

Cada MAR/OFSwitch tiene asociado un prefijo de red de 64 bits. En el caso de SR-DMM el prefijo es calculado por la aplicación desarrollada que se ejecuta sobre el controlador SDN. Por el contrario, en la implementación NB-DMM, el prefijo de red es calculado por el propio agente de movilidad.

En el caso de la solución NB-DMM, los distintos MAR y el CMD ejecutan un sistema operativo Linux con un *kernel* que ha sido preparado con unas características específicas, siendo necesario compilar el *kernel* añadiendo el soporte a la movilidad

IPv6, que no está activado por defecto. Por ello, es necesario iniciar un proceso de configuración del *kernel*, activando las características de movilidad necesaria para, posteriormente, compilarlo de nuevo y, así, los cambios surtan efecto. Las opciones relacionadas con el soporte de movilidad que tienen que ser habilitadas se muestran a continuación.

Linux Kernel (Características para el soporte a la movilidad)



Una vez llevado a cabo el proceso descrito anteriormente, se instala la implementación de NB-DMM en los nodos que actúan como MAR y CMD. Por el contrario, tanto el CN como el MN utilizan un sistema operativo Linux que no requiere ningún cambio en el *kernel* ni *software* adicional.

En el caso de SR-DMM no es necesario actualizar la pila de protocolos TCP/IPv6 de ningún nodo.

La evaluación experimental llevada a cabo se ha centrado en el análisis de la latencia de *handover* en ambas propuestas. La latencia de *handover* ha sido tomada como el intervalo de tiempo en el cual el MN pierde la conectividad IP tras realizar un proceso de *handover*.

En la evaluación experimental se ha usado Wireshark para extraer y analizar los eventos producidos durante la misma. El proceso que se ha seguido ha sido el siguiente: inicialmente, el MN se asocia al MAR1/OFSwitch1 y comienza a recibir tráfico UDP desde el CN. Posteriormente, el MN visita MAR2/OFSwitch2 y MAR3/OFSwitch3, manteniendo activo el flujo inicial asociado al primer agente de movilidad. Una vez hecho esto, el MN vuelve al MAR1/OFSwitch1 y repite el proceso, obteniéndose un total de 300 *handovers* para cada solución. La Figura 4.10 muestra los resultados obtenidos, representando la función de distribución acumulada (CDF) experimental para los valores de tiempo de recuperación de flujos UDP de ambas soluciones.

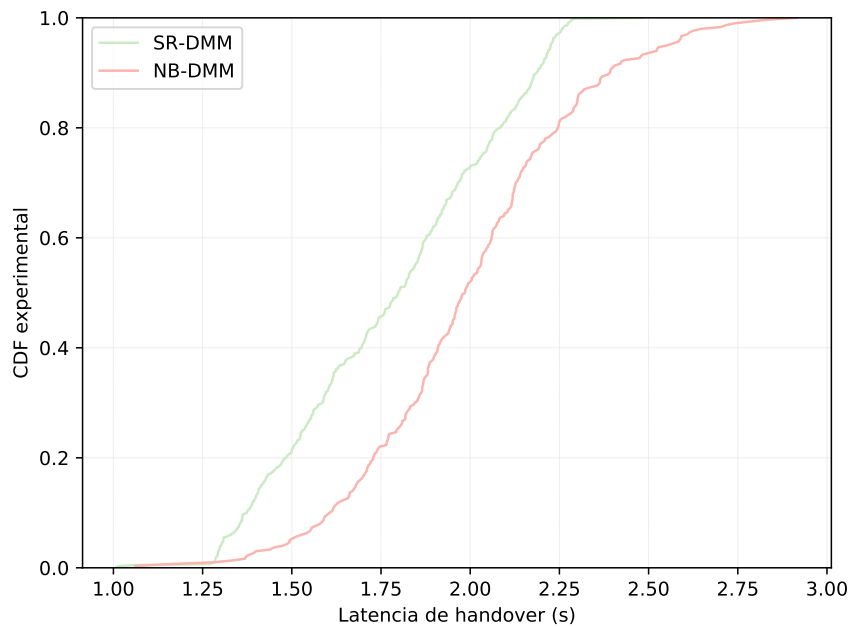


Figura 4.10: Función de distribución acumulada (CDF) experimental de las medidas de latencia *handover*

Además, la Tabla 4.2 muestra la media y desviación típica de los resultados obtenidos para ambas propuestas. Tal y como puede observarse, en general los resultados de latencia de *handover* para la solución SR-DMM son mejores que en NB-DMM.

Tabla 4.2: Media y desviación típica de los resultados obtenidos de latencia de *handover*

	Media (s)	Desviación típica (s)
NB-DMM	2.01	0.31
SR-DMM	1.78	0.29

También se ha llevado a cabo otro tipo de análisis consistente en la evaluación de los diferentes componentes que afectan a la latencia de *handover*, segregando el tiempo

total de desconexión en los diferentes niveles involucrados en la comunicación. Las pruebas se han realizado sobre el *tesbed* descrito en la Figura 4.9 y el tráfico utilizado en los experimentos ha sido UDP. Se han identificado tres eventos que afectan a las medidas de latencia de *handover*:

- L2: el *handover* de nivel 2 es el tiempo requerido para llevar a cabo el cambio desde un MAR/OFSwitch a otro. Este componente ha sido medido como el intervalo entre dos mensajes de control IEEE 802.11. Durante este tiempo, el MN pierde su enlace con el punto de acceso inalámbrico hasta establecerlo con el nuevo, enviando un RS a su nuevo ancla de movilidad.
- LSDN: es el tiempo transcurrido desde que el MN envía un RS al OFSwitch en la nueva red visitada, hasta que este le comunica al MN los parámetros IP de esa red con el envío de un RA. La solución NB-DMM no contempla este tiempo, debido a que no utiliza ningún tipo de mecanismo SDN.
- L3: es el tiempo transcurrido desde que el MN configura su nueva dirección IPv6, hasta la recepción del primer paquete perteneciente al flujo previamente iniciado.

La Figura 4.11 muestra en detalle los parámetros que afectan a la latencia de *handover* para la solución NB-DMM y la propuesta SR-DMM.

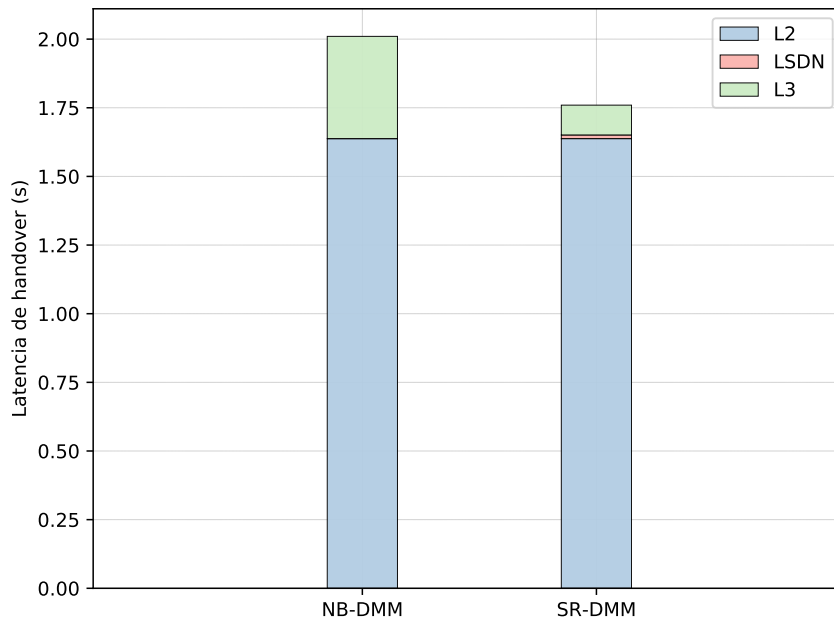


Figura 4.11: Latencia de *handover* segregada para NB-DMM y SR-DMM

Tal y como se puede observar, la mayor parte del tiempo es consumido en la desconexión del MN de la red inalámbrica previa y la posterior asociación a la nueva red (tiempo L2). La tecnología utilizada en este nivel es IEEE 802.11g. Por otro lado, el tiempo invertido por el controlador SDN para el correspondiente procesamiento (LSDN) es muy bajo, constituyendo un 0.3 % del total de la latencia de *handover*. NB-DMM no consume tiempo LSDN, ya que no requiere de mecanismos SDN. Por último, la recuperación de flujo (L3) depende en gran medida del tipo de solución: NB-DMM usa el 18 % del tiempo de *handover* y SR-DMM solamente el 6 %, dejando constancia de los beneficios aportados por la tecnología SDN.

4.5. Conclusiones

En este capítulo se presenta un enfoque novedoso de gestión de la movilidad distribuida basado en SDN, llamado SR-DMM, que permite una gestión eficiente del plano de datos. La tecnología SDN ofrece una separación natural del plano de control y el plano de datos para entornos 5G, donde la gestión de la movilidad distribuida se considera necesaria en los despliegues de arquitecturas de redes móviles de nueva generación. Por ello, se han utilizado las capacidades y beneficios que proporciona SDN, planteando una solución cuyo principal objetivo es aportar flexibilidad, escalabilidad y fiabilidad al proceso de movilidad en arquitecturas de red distribuidas.

En este capítulo, se ha llevado a cabo una evaluación de rendimiento analítica y experimental entre la propuesta SR-DMM y la solución más representativa de gestión de la movilidad distribuida (NB-DMM).

El modelo analítico revela los beneficios que aporta SR-DMM al rendimiento global de la red en términos de coste de señalización y coste de entrega de paquetes. Estas mejoras son conseguidas debido a que SR-DMM reduce la complejidad del plano de datos y la gestión de túneles IP sobre IP, evitando el uso de los mismos durante el movimiento de los usuarios móviles.

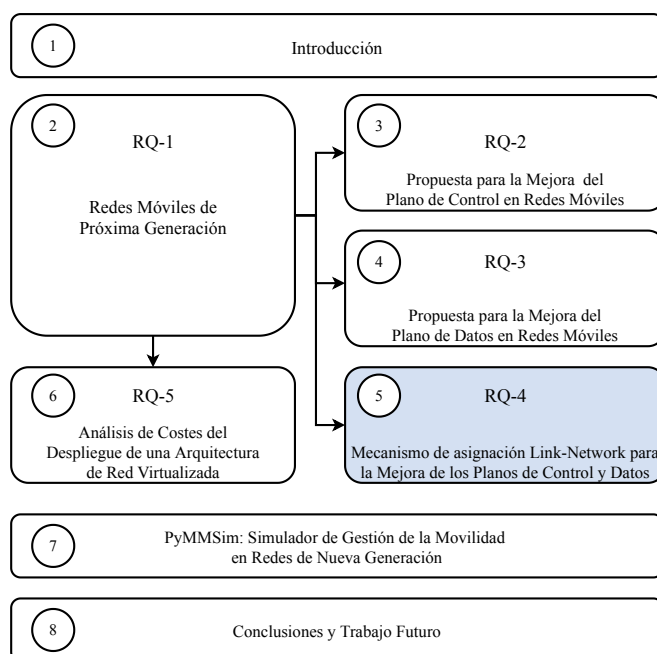
Por otro lado, para realizar la evaluación experimental, se ha llevado a cabo la implementación de nuestra propuesta sobre el controlador SDN Ryu, permitiendo así llevar a cabo la evaluación experimental, utilizando implementaciones reales de NB-DMM y nuestro mecanismo (SR-DMM) sobre un *testbed* de movilidad real. Los resultados demuestran que los mecanismos propios de SDN introducen una latencia mínima en el proceso de *handover*. También, se ha mostrado que la tecnología SDN puede reducir la complejidad de la gestión de la movilidad, convirtiéndose en un aspecto clave y a tener en cuenta en el diseño de redes móviles.

Por último, se puede concluir que SR-DMM es una solución de gestión de la movilidad que aporta flexibilidad a las redes tradicionales y aprovecha las capacidades de programabilidad disponibles en las redes de nueva generación. Además de esto, añade una mejora sustancial en el coste de entrega de paquetes (plano de datos) con respecto a la solución NB-DMM, que emplea mecanismos de tunelización para redirigir los flujos de datos hasta la ubicación del MN, añadiendo cierta complejidad al plano de datos.

Capítulo 5

Mecanismo de asignación Link-Network para la Mejora de los Planos de Control y de Datos

En este capítulo se desarrolla un mecanismo novedoso que tiene como objetivo la asignación eficiente de estaciones bases a los nodos de borde de la red de acceso, consiguiendo importantes mejoras de rendimiento en el plano de control y en el plano de datos de los protocolos de gestión de la movilidad IP.



Se ha optado por una estrategia de este tipo ya que, en general, las soluciones de gestión de la movilidad tienen como objetivo balancear la sobrecarga de señalización y los costes asociados al plano de datos. Tal y como se ha discutido en capítulos anteriores, la disminución de uno los costes asociados a un plano en concreto impacta negativamente en el otro, y viceversa.

En definitiva, el capítulo se organiza tal y como se presenta a continuación:

- La Sección 5.1 contextualiza el problema que afecta en la actualidad a las redes móviles, debido al incremento del tráfico de datos y de los nuevos casos de uso que han surgido con la aparición de la tecnología 5G, que traen asociados unos servicios y aplicaciones con estrictos requisitos de latencia y fiabilidad.
- La Sección 5.2 presenta la integración de los mecanismos de gestión de la movilidad expuestos en el Capítulo 2 sobre la arquitectura de la red de acceso propuesta por el 3GPP.
- En la Sección 5.3 se describe el modelado del sistema y se presenta la formulación de un problema de optimización para minimizar el impacto de la asignación de estaciones base a *routers* de acceso. Además, se desarrolla un nuevo algoritmo llamado *Link-Network Assignment* (LNA), con el objetivo de llevar a cabo una asignación adecuada mediante la recopilación de información de la topología de la red de acceso y el análisis de la distribución de las estaciones base en un área determinada.
- En la Sección 5.4 se definen las métricas relacionadas con los planos de control y de datos utilizadas para la evaluación de la estrategia propuesta.
- La Sección 5.5 presenta una evaluación de rendimiento del algoritmo propuesto LNA con respecto a otros algoritmos de referencia, en términos de costes de movilidad ampliamente utilizados, permitiendo evaluar el rendimiento global de la red móvil.
- La Sección 5.6 muestra las conclusiones que han sido identificadas a partir del análisis realizado de la evaluación llevada a cabo.

5.1. Introducción

En los últimos años, el mundo ha sido testigo de una evolución considerable de las comunicaciones móviles, debido a la gran proliferación de dispositivos móviles

e inteligentes que generan una cantidad de tráfico de datos sin precedentes. Según estudios recientes, se espera que el número de suscripciones móviles alcance los 8900 millones a finales de 2022 [19], produciéndose un crecimiento importante de dispositivos interconectados entre 2017 y 2022, incrementándose en un 37 %. Además, tal y como se ha mencionado anteriormente, se prevé que el tráfico global de datos móviles se incremente siete veces durante el periodo 2017-2022, estimándose un consumo mensual de 77.5 exabytes en 2022 [18, 19]. Además, las previsiones indican que en 2023 habrá más de mil millones de clientes haciendo uso de redes móviles 5G.

Como el número de usuarios móviles está creciendo a una velocidad vertiginosa, la nueva generación de comunicaciones móviles (5G) está evolucionando para hacer frente a este crecimiento y garantizar servicios y aplicaciones emergentes de acuerdo a las demandas específicas de los usuarios móviles; planteándose como una tecnología que dará respuesta a retos complejos y exigiendo un avance tecnológico significativo, ya que deben ser tenidas en cuenta aplicaciones y servicios con estrictos requisitos de rendimiento.

En entornos de red de próxima generación es complejo producir mejoras de alta fiabilidad y de baja latencia de forma simultánea. Por ello, nuevas tecnologías de vanguardia, como *Mobile Edge Computing* (MEC), permiten obtener mejoras en este sentido, pero es necesario ubicar las capacidades de computación cerca de los usuarios móviles [23], tal y como se explicó en el Capítulo 2.

Adicionalmente, 5G PPP (*5G Infrastructure Public Private Partnership*), a través de uno de sus grupos de trabajo (*Architecture Working Group*¹), ha sentado las bases de los casos de uso emergentes en las redes 5G [42], donde uno de los aspectos clave es el soporte a la movilidad. De esta manera, las redes móviles 5G imponen estrictos desafíos de gestión de la movilidad, ya que la mayoría de los servicios, aplicaciones y casos de uso contemplan rigurosas restricciones de rendimiento [43].

Los mecanismos de gestión de la movilidad proporcionan soporte a la movilidad de manera transparente en entornos de red heterogéneos, sin interrupciones a nivel de red, manteniendo las comunicaciones activas durante el movimiento de los usuarios móviles entre distintas redes de acceso. Estos protocolos se pueden clasificar en enfoques centralizados CMM y distribuidos DMM, tal y como se explicó en el Capítulo 2. Estos mecanismos de gestión de movilidad introducen sobrecarga de señalización cuando ofrecen soporte a la movilidad. Este aspecto, junto a la densificación celular producida por 5G, hará que aumente el tráfico total de señalización, degradando los requisitos de QoS y QoE. Por tanto, los operadores buscan soluciones innovadoras para la

¹<https://5g-ppp.eu/5g-ppp-work-groups/>

optimización de los procedimientos de gestión de la movilidad dentro de la arquitectura de red 5G. Los operadores de red móvil tendrán que tener en cuenta los aspectos mencionados anteriormente al planificar las arquitecturas de red de próxima generación, con el objetivo de mejorar el rendimiento de la red y diseñar redes para soportar los requisitos de los casos de uso emergentes en 5G. Nuevas líneas de investigación incluyen el uso de técnicas de *clustering* para la asignación de estaciones base a nodos de acceso y la formulación de un problema de optimización para mejorar el rendimiento global de la red [23]. Por lo tanto, la asignación óptima entre estaciones base y la red de acceso se ha convertido en un desafío que deben abordar los operadores de redes móviles de nueva generación.

5.2. Gestión de la movilidad en arquitecturas 3GPP

En esta sección se describe la integración de los protocolos de gestión de la movilidad, explicados en la Sección 2.3, sobre las arquitecturas de acceso propuestas por el 3GPP [66], proporcionando una visión general del soporte a la movilidad IP desde esta perspectiva.

En la actualidad, la mayoría de arquitecturas de red desplegadas cuentan con un número limitado de anclas centralizadas que administran el tráfico de miles de usuarios móviles, pero estos enfoques centralizados tienen ciertos problemas que se han identificado en numerosos trabajos de investigación [9, 69, 70].

Por ello, la evolución hacia los enfoques distribuidos ha revelado mejoras en relación al uso de los recursos de la red, proporcionando un nuevo concepto para proporcionar soporte a la movilidad, distribuyendo las funciones de movilidad de los planos de control y de datos entre las entidades de red ubicadas en el borde de la red de acceso.

A continuación se proporciona una visión general de cómo se realiza la integración sobre la red de acceso [120] de los enfoques más representativos de gestión de la movilidad (PMIPv6 y NB-DMM), explicados en detalle en la Sección 2.3.1 y Sección 2.3.2, respectivamente.

5.2.1. Gestión de la movilidad centralizada en redes de acceso 3GPP

El protocolo de gestión de movilidad centralizada más representativo es PMIPv6 [68], donde un único ancla de movilidad gestiona la señalización y el tráfico de datos de los nodos móviles. PMIPv6 ofrece soporte de movilidad basado en la red, es decir, se proporciona soporte para la movilidad a los nodos móviles sin necesidad de que

intervengan de manera activa en la señalización. Para ello, este protocolo gestiona la movilidad en redes de acceso 3GPP introduciendo una entidad de red llamada *Packet Data Network Gateway* (PGW) que actúa como *Local Mobility Anchor* (LMA). El PGW es responsable de gestionar las sesiones abiertas de cada MN. Además, PMIPv6 reubica ciertas funcionalidades desde el MN hacia la red, concretamente la detección del movimiento y las operaciones de señalización, que serán llevadas a cabo por una nueva entidad funcional llamada *Serving Gateway* (SGW). Esta entidad se suele utilizar en el borde de la red de acceso, ejerciendo de *router* de acceso (AR) para el MN, detectando los movimientos del mismo e informando al PGW de tales sucesos. La visión general de este protocolo desde la perspectiva de la red de acceso 3GPP se muestra en la Figura 5.1.

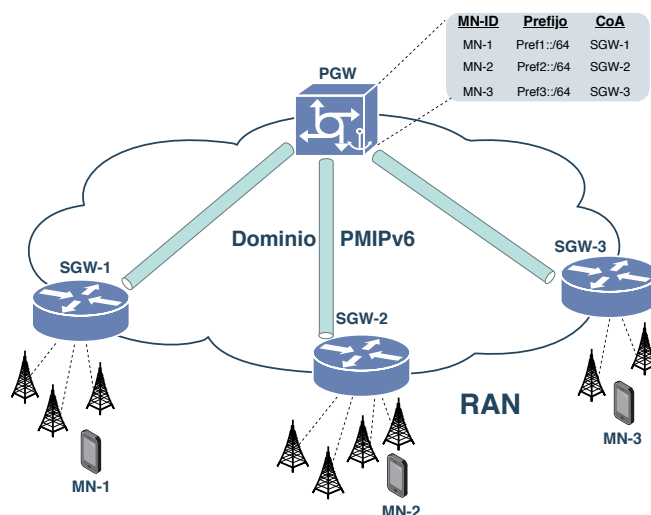


Figura 5.1: Visión general de la arquitectura de PMIPv6 sobre la red de acceso 3GPP

Cuando un MN realiza su movimiento por el dominio y cambia de SGW, este envía un mensaje *Proxy Binding Update* PBU al PGW para establecer un túnel IP sobre IP bidireccional entre el SGW y el PGW. Además, el PGW contesta al SGW con un mensaje de *Proxy Binding Acknowledgement* (PBA). En definitiva, se establecerán túneles entre el PGW y cada SGW por el que vaya pasando el MN; y gracias a ello, el PGW será capaz de capturar todo el tráfico IP dirigido al MN, pudiendo redireccionarlo correctamente hacia él mismo. Cada MN tiene configurada una *home address* con un prefijo de red gestionado por el PGW.

La operación de PMIPv6 muestra que los planos de control y de datos son altamente jerárquicos, siendo necesario el establecimiento de numerosos túneles IP sobre IP entre el PGW y cada SGW para proporcionar con eficacia soporte a la movilidad, causando problemas de rendimiento y escalabilidad.

5.2.2. Gestión de la movilidad distribuida en redes de acceso 3GPP

Tal y como se ha mencionado en reiteradas ocasiones, el enfoque más representativo de gestión de la movilidad distribuida es NB-DMM [78]. Con este protocolo, el MN se encuentra exento de participar en la señalización relativa a la gestión de su propio contexto de movilidad, al igual que ocurre con PMIPv6. De esta manera, no es necesaria la actualización del *software* de red para el soporte a la movilidad en el MN, ya que los agentes de movilidad distribuida llevarán a cabo las funciones de señalización requeridas. NB-DMM es una de las primeras propuestas diseñadas por el grupo de trabajo de DMM².

En las arquitecturas NB-DMM, las funcionalidades de gestión de la movilidad son distribuidas por el borde de la red de acceso, en unas entidades llamadas DMM *gateways* (DMM-GW), con el objetivo de obtener una red más plana separando los planos de datos y de control. Estos nuevos agentes de movilidad, además de proporcionar las funciones requeridas para anclar las sesiones de los nodos móviles a él, proporcionan funciones para la actualización del contexto de movilidad de todos los nodos móviles que él mismo gestiona. De esta manera, se elimina la jerarquía de los planos de control y de datos, ya que las funciones de las entidades SGW y PGW son llevadas a cabo por las entidades DMM-GW.

La visión general de este protocolo desde la perspectiva de la red de acceso 3GPP [120] se muestra en la Figura 5.2.

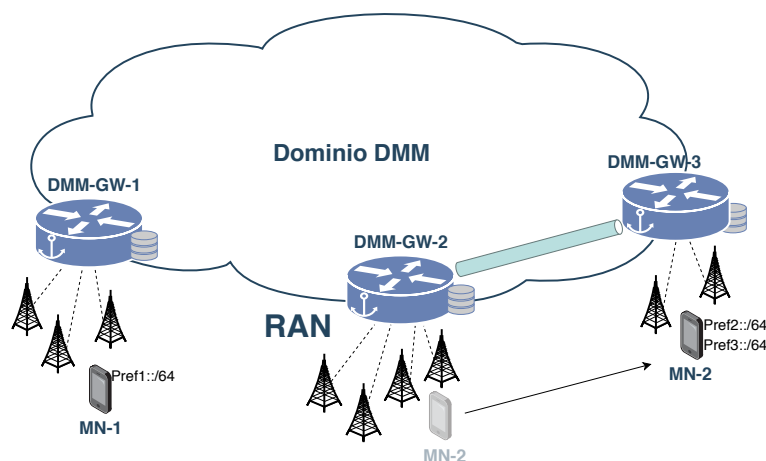


Figura 5.2: Visión general de la arquitectura de NB-DMM sobre la red de acceso 3GPP

²IETF DMM Working Group: <https://datatracker.ietf.org/wg/dmm/>

El DMM-GW detecta las conexiones de cada MN a la red de acceso y le proporciona un *Home Network Prefix* (HNP). Además, cuenta con una caché de vínculos local para almacenar la información de los nodos móviles que él mismo ha registrado en el dominio.

Por otro lado, para mantener una tabla de correspondencias general, se define otra entidad llamada *Context Mobility Database* (CMD). Se trata de una base de datos que proporciona el contexto de movilidad de todos los nodos móviles registrados en el dominio DMM.

Cuando se produce un *handover* antes del final de la sesión, el tráfico de datos se envía a través de un túnel establecido entre el DMM-GW actual y el DMM-GW donde fue anclada esa sesión inicialmente. En un *handover*, el nuevo DMM-GW recupera del CMD, a través de mensajes PBU y PBA, las direcciones IP de los DMM-GW que actúan como ancla para las sesiones en curso del MN. Posteriormente, se procede a la actualización de la ubicación del MN en cada uno de los DMM-GW que actúan como ancla para alguna sesión. Este proceso se lleva a cabo a través de mensajes de señalización PBU y PBA.

5.3. Modelado del sistema y formulación del problema

En esta sección, se presenta el modelo del sistema sobre el que se definirá el problema de optimización para minimizar el impacto de la asignación de estaciones base a los nodos del borde de la red de acceso, sin pérdida de rendimiento, en términos de coste de entrega de paquetes (C_{PD}) y coste de señalización (C_S) asociados a los protocolos de gestión de la movilidad.

Se considera una red de acceso representada como un grafo no dirigido $G = (V, E)$, donde V y E denotan los conjuntos de nodos y enlaces (aristas) respectivamente. Sea $K \subseteq V$ el conjunto de los *routers* de acceso que sirven y proporcionan acceso a los nodos móviles a través de un conjunto de estaciones base B , donde cada estación base es denotada por $b_i (1 \leq i \leq |B|)$. Este conjunto B proporciona cobertura total a un área geográfica determinada y cada ubicación viene dada por $\{L_{b_i}\}_{b_i \in B}$, donde $L_{b_i} \in \mathbb{R}^2$ representa el espacio bidimensional donde se ubicarán las estaciones base.

Por lo tanto, cada *router* de acceso $\{k_j\}_{j \in K}$ sirve a un subconjunto de estaciones base $B_k \subseteq B$ dentro de un dominio de red. Los *routers* de acceso son definidos como los nodos de primer salto que pueden ser considerados como el enlace entre el nivel físico y el nivel de red. Además, N denota el conjunto de nodos móviles que se mueven por el dominio de red y se conectan a las estaciones base $b_i \in B$; siendo cada MN definido por $N_j (1 \leq j \leq |N|)$.

Adicionalmente, se asume que cada estación base b_i es asociada a un *router* de acceso del dominio de movilidad, tal y como se muestra en la Figura 5.3, y cada *router* de acceso k_j gestiona un conjunto determinado de estaciones base.

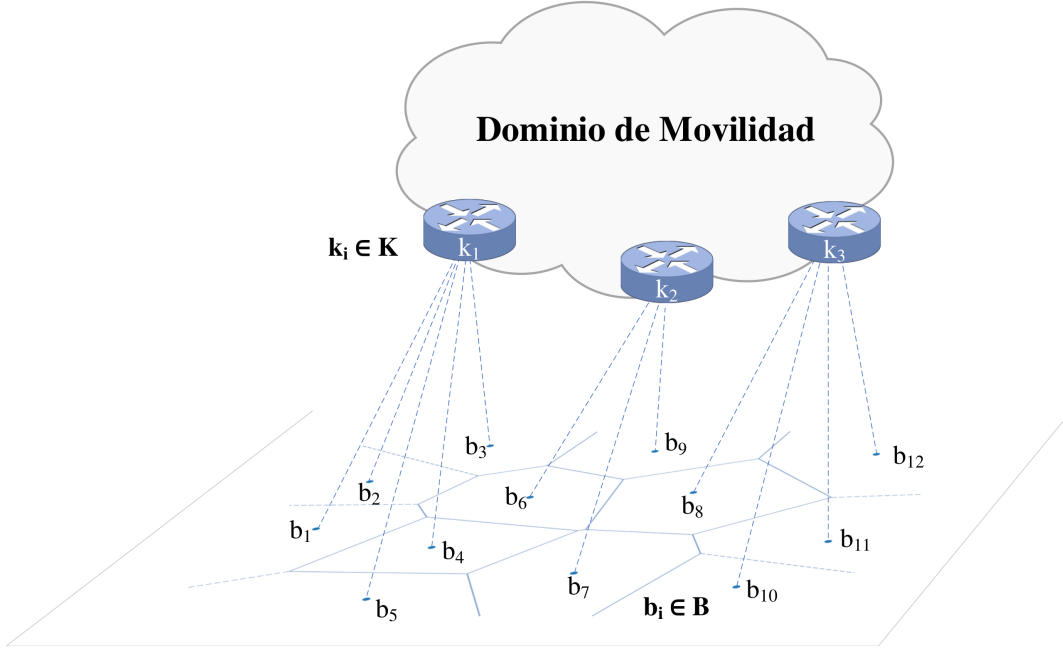


Figura 5.3: Asignación de las estaciones base al dominio de movilidad

Tomando como base la definición previa del modelo de red y su formalización, definimos una variable de decisión x_{mr}^{ps} tal y como se muestra a continuación:

$$x_{mr}^{ps} = \begin{cases} 1 & \text{si la estación base } m \text{ es asignada al } \textit{router} \text{ de acceso } r \\ & \text{y la estación base } p \text{ es asignada al } \textit{router} \text{ de acceso } s. \\ 0 & \text{en caso contrario.} \end{cases}$$

Teniendo en cuenta esta variable de decisión, la asignación óptima entre la red de acceso y las estaciones base es definida como un problema de optimización:

$$\text{Min}_{x_{mr}^{ps}} \quad F = \sum_{m \in B} \sum_{r \in K} \sum_{p \in B} \sum_{s \in K} TC x_{mr}^{ps} \quad (5.1)$$

sujeto a:

$$\sum_{r \in K} \sum_{p \in B} \sum_{s \in K} x_{mr}^{ps} + \sum_{m' \in B} \sum_{r' \in K} \sum_{s' \in K} x_{m'r'}^{p's'} = 1, \quad (5.2)$$

$$\forall m = p' \in B, m = p' = 1, \dots, B$$

$$\sum_{m \in B} \sum_{p \in B} \sum_{s \in K} x_{mr}^{ps} + \sum_{m' \in B} \sum_{r' \in K} \sum_{p' \in B} x_{m'r'}^{p's'} \leq th_j, \quad (5.3)$$

$$\forall r = s' = j \in K, r = s' = j = 1, \dots, K$$

$$x_{mr}^{ps} \in \{0, 1\}, \forall m \in B, r \in K, p \in B, s \in K \quad (5.4)$$

Para cada asignación, se define un coste total TC como la suma de dos parámetros críticos relacionados con los protocolos de gestión de la movilidad: coste de señalización (C_S) y coste de entrega de paquetes (C_{PD}).

$$TC = C_S + C_{PD} \quad (5.5)$$

C_S considera la carga de tráfico generada por los mensajes de señalización cuando se produce un proceso de *handover* de nivel tres, con el objetivo de mantener las sesiones activas de cada nodo móvil (MN). Por otro lado, C_{PD} depende del tamaño de los mensajes de datos multiplicado por el número de saltos necesarios para reenviar los paquetes al MN. La Restricción 5.2 indica que una estación base ($m = p' \in B$) es asignada a un único *router* de acceso y la Restricción 5.3 está relacionada con el balanceo de estaciones base entre los diferentes *routers* de acceso. Se asume que un determinado *router* de acceso ($r = s' \in K$) no puede servir a más de un número específico de estaciones base, que viene determinado por un *threshold* (th_j).

El modelo anterior calcula la asignación óptima entre el nivel físico y la red de acceso. En este contexto, el tamaño del problema depende del número de estaciones base y los nodos de la red de acceso. Cuando se reduce el tamaño del problema, por ejemplo, 15 estaciones base y 3 *routers*, se encuentra una solución óptima en menos de un segundo. Sin embargo, si el número de estaciones base aumenta o se utilizan topologías de red a gran escala, la complejidad de computación aumenta exponencialmente. Por esta razón, se propone una nueva estrategia para resolver este problema en tiempo polinomial. El Algoritmo 4 define el enfoque *Link-Network Assignment* (LNA) detalladamente.

Algoritmo 4 *Link-Network Assignment* (LNA).

Entrada: Conjunto de estaciones base $B = \{b_1, \dots, b_{|B|}\}$, conjunto de *routers* de acceso $K = \{k_1, \dots, k_{|K|}\}$ y la topología de red NT

Salida: Diccionario *DictAssoc* con la asociación final entre los *clústeres* de estaciones base (centroides) y los *routers* de acceso

▷ **Primera fase**

1: $C = \text{k-means++}(B, |K|)$;

▷ **Segunda fase**

2: $D_{CKM} = \text{getEuclideanDistances}(C)$;

3: $D_{AR} = \text{allShortestPath}(NT, K)$;

▷ **Tercera fase**

4: **for** $c_i \in C$ **do**

5: $M_C[i] = \sqrt{c_{ix}^2 + c_{iy}^2}$;

6: **end for**

7: **for** $k_i \in K$ **do**

8: $CC_{AR}[j] = |K| / \sum_{k_j \in K} D_{AR}[k_i][k_j]$;

9: **end for**

▷ **Cuarta fase**

10: $MC = \text{argmax}(M_C)$; $cc = \text{argmin}(CC_{AR})$

11: **for** $j = 1$ to $|K|$ **do**

12: $d_c[j] = D_{CKM}[MC][j]$;

13: $d_{ar}[j] = D_{AR}[cc][j]$;

14: **end for**

15: $d_c = \text{sorted}(d_c)$; $d_{ar} = \text{sorted}(d_{ar})$

16: **for** $i = 1$ to $|K|$ **do**

17: $\text{DictAssoc}\{\text{"}d_c[i]\text{"}\} = d_{ar}[i]$;

18: **end for**

19: **return** *DictAssoc*;

Este algoritmo recopila la información de la topología de la red de acceso y examina la distribución de las estaciones base para realizar una asignación adecuada, y está compuesto por cuatro fases principalmente, que son definidas a continuación:

- ▷ **Primera fase:** Un conjunto de observaciones de datos (conjunto de estaciones base $B = \{b_1, \dots, b_{|B|}\}$) se clasifica en un número específico $|K|$ de agrupaciones (*clústeres*), que coinciden con el número de *routers* de acceso de la topología de red, utilizando un algoritmo no supervisado, ampliamente utilizado, llamado k-means++ [121]. Este algoritmo presenta una ventaja importante con respecto a la inicialización de los centroides, con el objetivo de mejorar el tiempo

computacional a la hora de realizar los cálculos, que puede volverse exponencial si se usa el algoritmo k-means original. Por lo tanto, k-means++ minimiza la distancia entre observaciones y centroides utilizando la distancia Euclídea, incluyendo un método de inicialización basado en una variable aleatoria uniforme; obteniendo, así, un conjunto adecuado de estaciones base al realizar la inicialización del algoritmo.

- ▷ **Segunda fase:** Para cada centroide c_i , se calculan las distancias $D_{CKM}(i, j)$ entre c_i y todos los demás ($c_j \in C$) usando la distancia Euclídea. Posteriormente, se construye la matriz de distancias D_{AR} de los nodos de acceso de la topología de red: para cada *router* de acceso k_i , se computa la distancia entre él mismo y los demás ($k_j \in K$) usando el algoritmo de Dijkstra.
- ▷ **Tercera fase:** Nuestro algoritmo selecciona el centroide con el valor del módulo más alto. Cada centroide c_i en un espacio bidimensional es definido como $c_i = [c_{ix}, c_{iy}]$. Por lo tanto, la norma Euclídea (módulo) de cada centroide puede ser calculada tal y como se define en la Ecuación 5.6, siendo c_{id} el valor de cada componente de un centroide determinado y D el número total de componentes de cada centroide, que en este caso serán dos, debido a que se está trabajando sobre un espacio bidimensional.

$$\|c\| = \sqrt{\sum_{d=1}^D |c_{id}|^2} \quad (5.6)$$

Además, para cada *router* de acceso $k_i \in K$, se calcula su centralidad (*closeness centrality*) [122] de acuerdo a la Ecuación 5.7, con el objetivo de seleccionar el nodo con el mínimo valor.

$$CC_{AR}(k_i) = \frac{|K|}{\sum_{k_j \in K} D_{AR}(k_i, k_j)} \quad (5.7)$$

Por tanto, CC_{AR} es un vector que almacena la centralidad (*closeness centrality*) para cada *router* de acceso.

- ▷ **Cuarta fase:** Posteriormente, el algoritmo realiza la primera asociación entre el centroide con mayor valor de módulo MC y el *router* de acceso con mínimo valor de *closeness centrality*, cc , el cual es obtenido del vector CC_{AR} , que fue calculado en la tercera fase de este algoritmo. Una vez hecho esto, se obtienen los vectores distancia d_c y d_{ar} . En primer lugar, d_c indica la distancia Euclídea entre MC y todos los demás centroides. En segundo lugar, d_{ar} almacena las distancias entre cc

(*router* de acceso con mínimo valor de *closeness centrality*) y los demás *routers*. Finalmente, ambos vectores son ordenados ascendentemente y la asociación final es realizada.

La Figura 5.4 muestra un ejemplo de la operación de la estrategia propuesta. En este caso, la red de acceso está formada por tres *routers* de acceso, por lo tanto, las estaciones base son agrupadas en tres *clústeres* utilizando el algoritmo k-means++. Posteriormente, se lleva a cabo la asociación final usando el Algoritmo 4 propuesto, tal y como se ha descrito anteriormente.

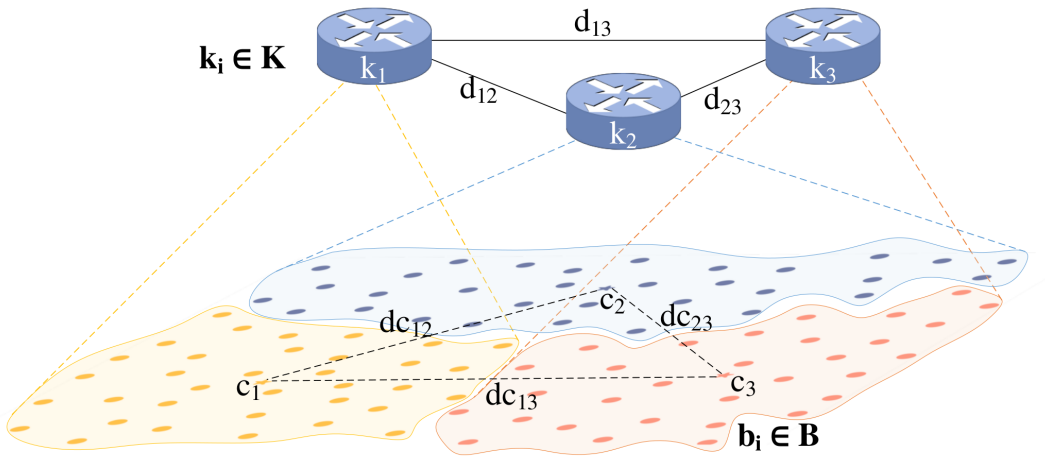


Figura 5.4: Ejemplo del algoritmo propuesto LNA para tres *routers* de acceso

Se ha realizado un análisis de complejidad computacional del algoritmo propuesto LNA. La Figura 5.5 muestra los tiempos de ejecución obtenidos en la evaluación de complejidad usando la infraestructura del supercomputador LUSITANIA II³, con 80 GB RAM y CPU Intel Haswell x86 (20 cores) a 2.6 GHz. El Apéndice A describe en detalle la infraestructura utilizada en todas las pruebas realizadas.

En esta evaluación, cada prueba se repite 20 veces para mejorar la precisión de los resultados con un intervalo de confianza del 95 %. Los resultados confirman que el tiempo de cálculo aumenta de manera lineal con respecto al número de estaciones base para nuestro algoritmo LNA. Sin embargo, el método de asignación óptimo muestra una tendencia exponencial.

³Características técnicas de LUSITANIA II:
<http://www.cenits.es/en/cenits/lusitania-II/lusitania-ii-specifications>

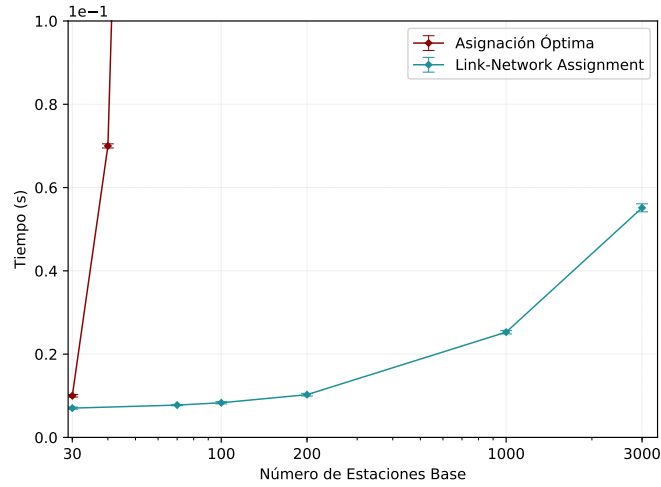


Figura 5.5: Evaluación de complejidad de computación del algoritmo propuesto LNA

5.4. Métricas de evaluación

En esta sección, se explican las métricas de rendimiento para evaluar la eficiencia de los protocolos de gestión de la movilidad desde dos puntos de vista: plano de control y plano de datos. El rendimiento del plano de control es evaluado a través de la sobrecarga introducida por el protocolo de movilidad en cuestión. Por otro lado, en el caso del plano de datos, su rendimiento es evaluado en términos de coste de entrega de paquetes. Estas métricas son ampliamente utilizadas en los análisis y evaluaciones de soluciones de movilidad [78, 85, 88, 93, 101], permitiendo determinar el rendimiento global de la red móvil.

Los protocolos de gestión de movilidad utilizados para la evaluación del rendimiento serán PMIPv6 y NB-DMM, como soluciones CMM y DMM, respectivamente. Los detalles y operación de cada uno de ellos se definieron en la Sección 5.2.

5.4.1. Plano de control

Una de las principales funcionalidades para cualquier protocolo de gestión de movilidad IP es el proceso de asegurar que la sesión de movilidad de cada MN se mantenga activa y actualizada mientras él realiza su movimiento por el dominio de red. Esto requiere de mensajes de control que deben enviarse entre los agentes de movilidad en la red.

Con el objetivo de evaluar el plano de control, una métrica relevante es el coste total de señalización relativo a la actualización de vínculos de movilidad durante una sesión

(C_S), el cual representa la carga de tráfico acumulada en el intercambio de mensajes de señalización. Este coste depende del tamaño de los propios mensajes de señalización y del número de *handovers* de nivel 3 realizados durante el intervalo de tiempo en el que la comunicación en el MN se mantiene activa.

En CMM (C_S^{CMM}), es necesaria la actualización de los vínculos de movilidad en el agente de movilidad (PGW). Por otro lado, en DMM (C_S^{DMM}), el DMM-GW, que está sirviendo al MN en un momento determinado (S_{GW}), recopila información de los DMM-GW previos y establece túneles bidireccionales IP sobre IP con ellos.

Por lo tanto, el coste de señalización durante el movimiento del MN para ambas soluciones (DMM y CMM) se resumen en las siguientes expresiones:

$$C_S^{CMM} = 2s_u h_{SGW-PGW} \quad (5.8)$$

$$C_S^{DMM} = 2s_u + 2s_u \sum_{i=1}^{n-1} h_{GW_i-S_{GW}} \quad (5.9)$$

donde n define el número de DMM-GW previos que establecen un túnel con el DMM-GW actual en la solución DMM. Además, h_{x-y} representa la distancia en número de saltos desde el nodo x al nodo y en la red y s_u indica el tamaño medio de los mensajes de señalización.

5.4.2. Plano de datos

Con respecto al plano de datos, una de las métricas que tienen un mayor impacto sobre el rendimiento global de la red es el coste de entrega de paquetes (C_{PD}). Aparte de la señalización relacionada con el proceso de movilidad, los paquetes de datos tienen que ser enviados desde el CN (*Correspondent Node*) al MN, y viceversa. Este valor está influenciado por el tamaño medio de los mensajes de datos multiplicado por el número de saltos necesarios para reenviar los paquetes desde el CN hasta el MN.

En las soluciones CMM, los paquetes se envían al usuario móvil utilizando un túnel IP sobre IP que encapsula los paquetes de datos entre el PGW y SGW. En protocolos DMM, cuando se produce un *handover*, el tráfico dirigido a la nueva ubicación del MN se enviará directamente hacia él, mientras que los paquetes asociados a sesiones abiertas

con anterioridad tendrán que ser encapsulados y enviados a través de un túnel entre el DMM-GW que actúa como ancla y el DMM-GW actual.

Por lo tanto, las expresiones que representan este coste para ambas soluciones se muestran a continuación:

$$C_{PD}^{MMM} = (s_d h_{CN-PGW} + (s_t + s_d) h_{PGW-SGW} + s_d h_{SGW-MN}) N_{p/s} \quad (5.10)$$

$$C_{PD}^{DMM} = (P_n C_{PD}^d + P_h C_{PD}^i) N_{p/s} \quad (5.11)$$

donde $N_{p/s}$ representa la tasa de transmisión de paquetes por flujo activo y s_d es el tamaño medio de los paquetes de datos. s_t es el tamaño medio de la cabecera de tunelización IPv6. Además, P_n y P_h son, respectivamente, las probabilidades de que un flujo de tráfico sea nuevo o de que un flujo siga abierto después de realizar un *handover*. Por lo tanto, C_{PD}^d y C_{PD}^i representan los costes de entrega de un paquete para los modos de funcionamiento directo e indirecto de DMM, respectivamente. Estos costes son expresados tal y como se muestra a continuación:

$$C_{PD}^d = s_d h_{CN-SGW} + s_d h_{SGW-MN} \quad (5.12)$$

$$C_{PD}^i = s_d h_{CN-GW} + (s_t + s_d) h_{GW-SGW} + s_d h_{SGW-MN} \quad (5.13)$$

5.5. Evaluación de rendimiento

Esta sección tiene como objetivo llevar a cabo un análisis del impacto de los costes de movilidad sobre el rendimiento general de la red, así como la evaluación del algoritmo propuesto LNA utilizando diferentes topologías de red. Estas topologías se basan en las topologías de red utilizadas en [123], variando el número de *routers* de acceso ($K = \{2, 4, 6, 8\}$). La Figura 5.6 muestra la topología de red con ocho *routers* de acceso (R10-R17). Las demás topologías ($K = \{2, 4, 6\}$) usadas en las simulaciones son

construidas a partir de esta. Por ejemplo, para $K = 2$, se eliminan los nodos desde R11 hasta R16 del nivel de acceso. De la misma manera, para $K = 4$, la topología está formada por todos los nodos desde R1 hasta R11, R16 y R17. Finalmente, para $K = 6$, se eliminan de la topología los nodos R13 y R14. Hay que tener en cuenta que cada nodo de la red juega un determinado papel, tal y como se muestra en la Tabla 5.1.

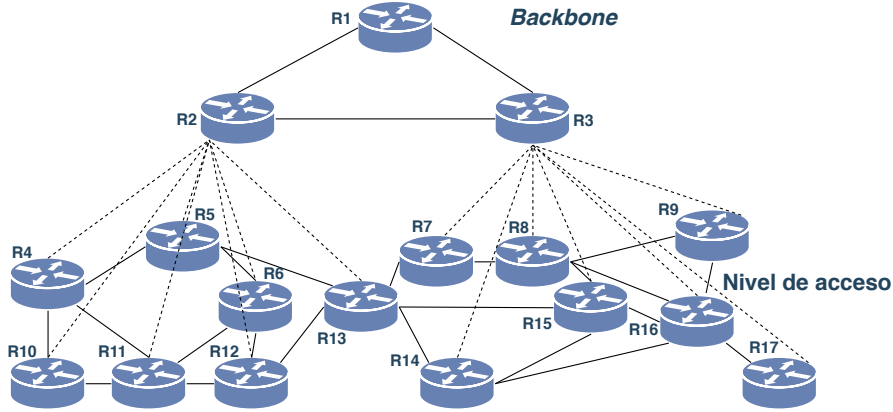


Figura 5.6: Topología de red utilizada en las simulaciones con ocho routers de acceso ($K = 8$)

Tabla 5.1: Funciones de los nodos de la topología de red para $K = 8$

	R1	R2-R9	R10-R17
Enfoque CMM	PGW	Router	SGW
Enfoque DMM	Router	Router	DMM-GW

Se han seleccionado varias topologías para proporcionar resultados con mayor fiabilidad, realizando pruebas con protocolos centralizados y distribuidos. A continuación, se presentan los parámetros de tráfico y movilidad utilizados en las simulaciones, así como los resultados numéricos obtenidos de los costes de movilidad.

Se ejecuta una simulación de Montecarlo de 500 iteraciones, proporcionando los valores medios y mejorando la precisión de los resultados con un intervalo de confianza del 95 %. El algoritmo de asociación propuesto LNA se evalúa mediante simulaciones utilizando Python con las bibliotecas NetworkX y SciPy [124], entre otras.

El escenario de simulación es una región cuadrada de 10×10 km², donde las estaciones base están distribuidas siguiendo un proceso llamado *Poisson Point Process* (PPP), cuya intensidad (λ_{BS}) coincide con el número medio de estaciones base (N_{BS}) por unidad de área (A) [125] y es obtenido como $\lambda_{BS} = N_{BS}/A$. Además, el área de cobertura de las estaciones base es modelada como una teselación de Poisson-Voronoi

[126], en un espacio bidimensional donde cada usuario móvil se conecta a la estación base más cercana.

La movilidad de usuario es definida a través del modelo de movilidad Random Waypoint [127] con una velocidad uniformemente distribuida entre 1 y 20 m/s . Cada simulación está formada por 200 usuarios que realizan su movimiento por el dominio de movilidad, conectándose a diferentes estaciones base. Estos usuarios móviles gestionan un conjunto de sesiones durante el tiempo de simulación, asumiéndose que el número de sesiones entrantes por usuario móvil sigue un proceso de Poisson con una tasa media $\lambda = 0.01$, y la duración de una sesión está exponencialmente distribuida con parámetro $\mu = 10$ [128]. También, se debe tener en cuenta que el flujo o tasa de datos de una demanda varía entre 1500 Kbps y 10 Mbps (por ejemplo, *stream* de vídeo) [129].

El rendimiento de la propuesta LNA es evaluada sobre este escenario de red, calculando los costes de movilidad: coste de señalización (C_S) y coste de entrega de paquetes (C_{PD}). En primer lugar, C_S se refiere a la sobrecarga de tráfico de señalización acumulada e introducida por el protocolo de gestión de la movilidad, cuando se produce un *handover* de nivel tres. En segundo lugar, el C_{PD} para una sesión está influenciado por el tamaño de los paquetes de datos multiplicado por el número de saltos necesarios para dirigir el tráfico de datos al usuario móvil.

Además, estos costes de movilidad son estudiados usando diferentes protocolos de gestión de la movilidad. Las redes móviles de próxima generación se pueden considerar idóneas para solventar los problemas de escalabilidad y confiabilidad que afectaban a las redes tradicionales, pero es necesario reducir la sobrecarga de señalización y producir mejoras en el plano de datos. Por lo tanto, en esta evaluación, la escalabilidad y fiabilidad es medida en términos de costes de señalización y costes de entrega de paquetes. Para analizar cómo los diferentes algoritmos de asignación afectan al rendimiento del plano de control, hemos realizado un conjunto de simulaciones sobre diferentes topologías de red ($K = \{2, 4, 6, 8\}$).

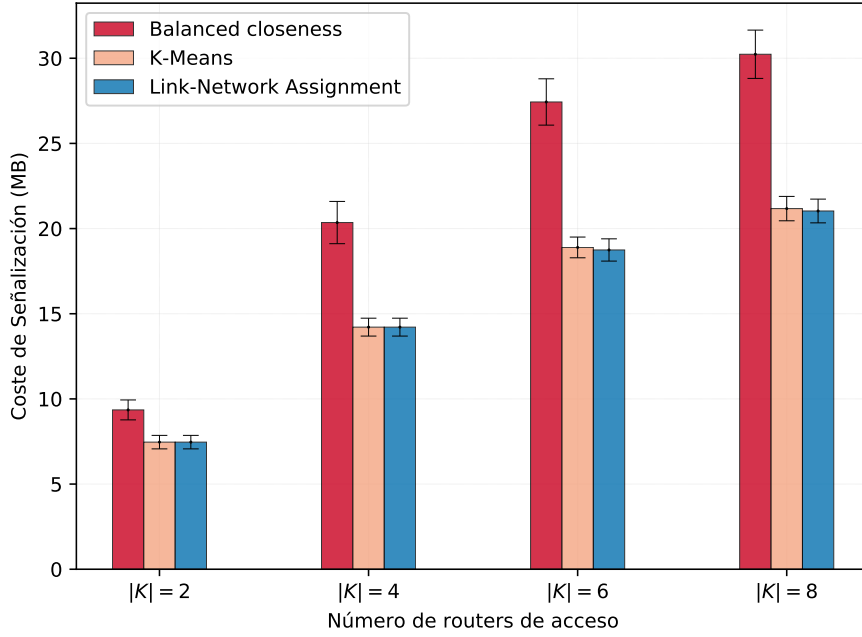


Figura 5.7: Evaluación del plano de control en protocolos CMM usando diferentes algoritmos de asignación

La Figura 5.7 y la Figura 5.8 muestran el coste de señalización acumulado para todas las conexiones generadas durante la simulación, proporcionando una comparativa de esta métrica en función del número de *routers* de acceso, para protocolos CMM y DMM, respectivamente. En estas evaluaciones, el rendimiento del algoritmo propuesto LNA, cuya operación fue descrita en la Sección 5.3, obtiene mejores resultados comparado con otros enfoques como:

- Algoritmo K-Means: divide el conjunto de estaciones base en $|K|$ *clústeres* y asocia dichos *clústeres* a los *routers* de acceso de manera aleatoria.
- Enfoque *Balanced Closeness*: selecciona K estaciones base y calcula las N estaciones base más cercanas a ella, siendo $N = |B|/|K|$. Posteriormente, el algoritmo construye $|K|$ grupos que son asociados a los *routers* de acceso.

En los enfoques DMM los agentes de movilidad son ubicados de manera distribuida por el borde de la red de acceso, gestionando todo el tráfico de señalización de los usuarios móviles. Debido a esta distribución de los nodos, DMM reduce las limitaciones y problemas asociados al tráfico de señalización que afectan a los enfoques CMM, proporcionando mejoras de rendimiento en el plano de control, tal y como muestra la Figura 5.8. Además, el coste de señalización es directamente proporcional al valor $|K|$,

para ambos protocolos (CMM y DMM). Por lo tanto, con respecto a los protocolos de gestión de movilidad, las soluciones CMM y DMM demuestran una tendencia clara a medida que aumenta el número de *routers* de acceso. Tal y como muestra la Figura 5.8, se observan mejoras relevantes en el coste de señalización, cuando se usa nuestro algoritmo de asociación (LNA) sobre arquitecturas de red distribuida.

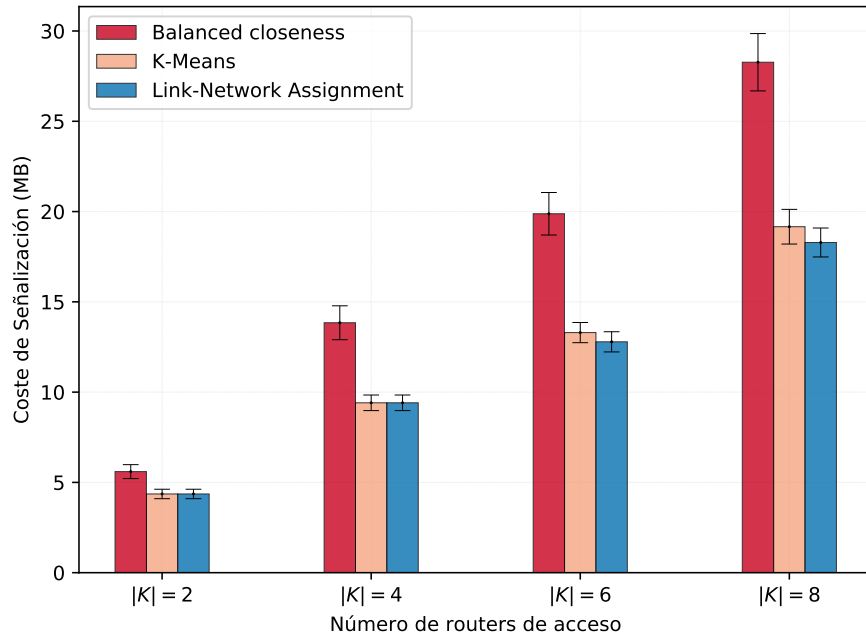


Figura 5.8: Evaluación del plano de control en protocolos DMM usando diferentes algoritmos de asignación

Con respecto al plano de datos, tal y como muestran las Figuras 5.9 y 5.10, el rendimiento es mejorado cuando se usa el enfoque propuesto LNA. Nuestro algoritmo presenta un menor impacto sobre el plano de datos, independientemente del protocolo utilizado, aunque proporciona más beneficios en soluciones CMM. Con respecto al protocolo de movilidad, ambos enfoques presentan un comportamiento similar, pero en el caso de CMM se revelan los problemas del encaminamiento subóptimo, explicados en la Sección 2.3.2, traducándose en un C_{PD} mayor. Otro aspecto influyente, y no menos importante, es que los valores de C_{PD} en DMM son menores que en CMM, debido también a la ubicación distribuida de los agentes de movilidad. Esta gestión eficiente de los recursos de red debido a la inclusión del algoritmo LNA puede facilitar el despliegue de arquitecturas de red móvil de nueva generación.

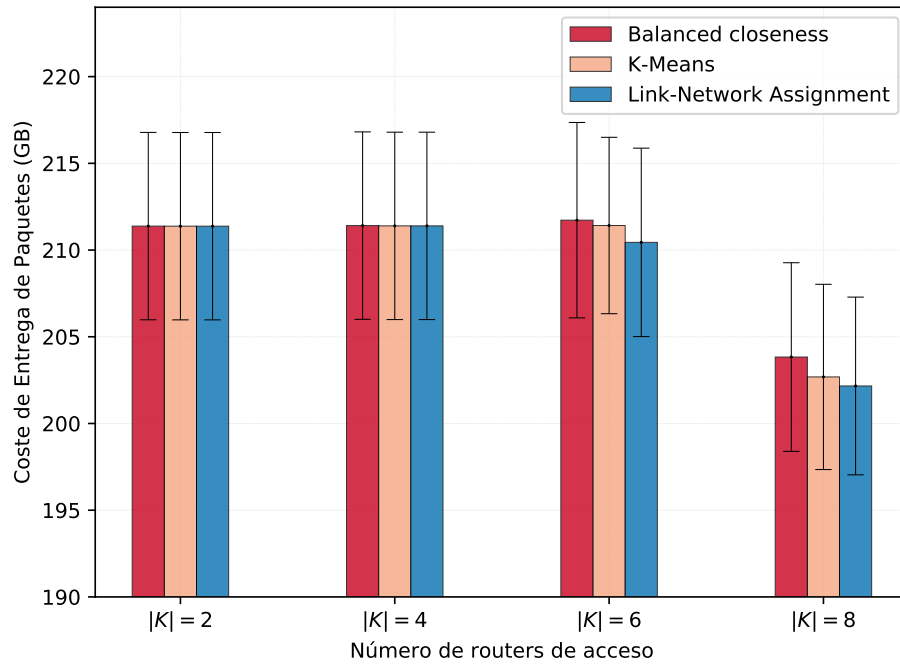


Figura 5.9: Evaluación del plano de datos en protocolos CMM usando diferentes algoritmos de asignación

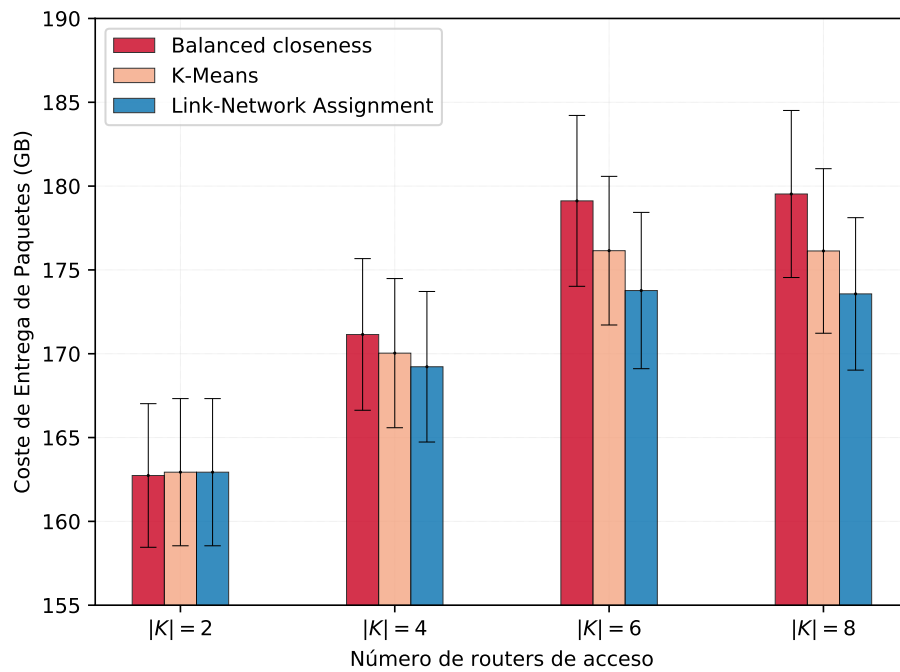


Figura 5.10: Evaluación del plano de datos en protocolos DMM usando diferentes algoritmos de asignación

Como se muestra en la Figura 5.9, si se incrementa el número de *routers* de acceso y se utilizan algoritmos de asignación basados en técnicas de *clustering* (K-Means y LNA), el C_{PD} disminuye con respecto a valores menores de $|K|$. Sin embargo, tal y como se muestra en la Figura 5.10, los costes asociados al plano de datos en el enfoque de gestión de la movilidad distribuida aumentan, excepto cuando $|K| = 8$. En este caso, el C_{PD} disminuye con respecto a $|K| = 6$ para ambos mecanismos de *clustering*: asignaciones K-Means y LNA.

Adicionalmente, el coste de entrega de paquetes está influenciado por los mecanismos de tunelización utilizados, produciendo diferencias significativas entre los enfoques centralizados y distribuidos. En los protocolos DMM esta tunelización es menos relevante, pero en las soluciones CMM producen una sobrecarga importante en la red de acceso. Por lo tanto, tal y como se muestra en la Figura 5.9, DMM logra ganancias significativas en términos de coste de entrega de paquetes, proporcionando mejoras sustanciales en el plano de datos. Este efecto es mayor cuando nuestro algoritmo LNA es utilizado para la asociación entre las estaciones base y la red de acceso.

Todos los resultados numéricos obtenidos son resumidos en la Tabla 5.2, la cual refleja la media y el error de los costes acumulados durante las simulaciones llevadas a cabo. Como se definió en la Sección 5.3, el coste total de movilidad es definido como la suma de dos componentes críticos relacionados con los protocolos de gestión de la movilidad. Estos parámetros son asociados al plano de control y al plano de datos. Los resultados demuestran que nuestro algoritmo de asignación tiene un menor impacto sobre el coste total de movilidad.

Si nos centramos en el plano de datos, los resultados muestran una reducción importante en el coste de entrega de paquetes. Sin embargo, vale la pena señalar que, aunque la métrica del plano de control puede considerarse insignificante con respecto al plano de datos, el plano de control se ha convertido en un factor crítico para los operadores de red cuando diseñan y planifican el despliegue de la red [130]. Además, considerando que la baja latencia representa uno de los mayores desafíos en las redes móviles de próxima generación, la reducción de la sobrecarga de señalización puede reducir la latencia global extremo a extremo.

Tabla 5.2: Evaluación de los protocolos de gestión de la movilidad usando diferentes algoritmos de asignación entre las estaciones base y los routers de acceso.

PROTOCOLO DE GESTIÓN DE LA MOVILIDAD CENTRALIZADA (PMIPv6)						
Análisis de algoritmos		$ K = 2$	$ K = 4$	$ K = 6$	$ K = 8$	
Balanced Closeness	$C_S(MB)$	9.35 ± 0.59	20.35 ± 1.24	27.43 ± 1.36	30.23 ± 1.42	
	$C_{PD}(MB)$	$211\,380.49 \pm 5401.64$	$211\,409.27 \pm 5403.60$	$211\,722.41 \pm 5632.99$	$203\,830.82 \pm 5437.97$	
	$C_S(MB)$	7.46 ± 0.40	14.21 ± 0.52	18.89 ± 0.61	21.17 ± 0.72	
K-Means	$C_{PD}(MB)$	$211\,375.00 \pm 5401.83$	$211\,393.59 \pm 5402.16$	$211\,415.46 \pm 5085.15$	$202\,684.63 \pm 5340.49$	
	$C_S(MB)$	7.46 ± 0.40	14.21 ± 0.52	18.74 ± 0.65	21.03 ± 0.70	
Link-Network Assignment	$C_{PD}(MB)$	$211\,375.93 \pm 5401.83$	$211\,393.59 \pm 5402.16$	$210\,442.95 \pm 5432.32$	$202\,162.96 \pm 5123.51$	
PROTOCOLO DE GESTIÓN DE LA MOVILIDAD DISTRIBUIDA (NB-DMM)						
Análisis de algoritmos		$ K = 2$	$ K = 4$	$ K = 6$	$ K = 8$	
Balanced Closeness	$C_S(MB)$	5.60 ± 0.39	13.84 ± 0.94	19.88 ± 1.18	28.27 ± 1.59	
	$C_{PD}(MB)$	$162\,737.46 \pm 4284.03$	$171\,150.97 \pm 4523.40$	$179\,118.25 \pm 5096.63$	$179\,529.07 \pm 4982.56$	
	$C_S(MB)$	4.36 ± 0.26	9.41 ± 0.43	13.30 ± 0.56	19.16 ± 0.96	
K-Means	$C_{PD}(MB)$	$162\,933.91 \pm 4390.05$	$170\,035.35 \pm 4449.34$	$176\,147.59 \pm 4432.98$	$176\,130.50 \pm 4908.59$	
	$C_S(MB)$	4.36 ± 0.26	9.41 ± 0.43	12.79 ± 0.56	18.29 ± 0.80	
Link-Network Assignment	$C_{PD}(MB)$	$162\,933.91 \pm 4390.05$	$169\,223.20 \pm 4490.39$	$173\,769.60 \pm 4661.60$	$173\,569.57 \pm 4547.01$	

5.6. Conclusiones

En este capítulo se propone una estrategia que realiza la asignación entre las estaciones base y la red de acceso de manera eficiente, llamada LNA. En primer lugar, se ha formulado un problema de optimización para minimizar el impacto de la asignación de estaciones base a los *routers* de acceso. Además, se ha desarrollado un nuevo algoritmo (LNA), con el objetivo de llevar a cabo una asignación adecuada mediante la recopilación de información de la topología de la red de acceso y el análisis de la distribución de las estaciones base en un área determinada. Finalmente, se lleva a cabo una evaluación de rendimiento del algoritmo propuesto con respecto a otros algoritmos de asignación, en términos de costes de señalización y costes de entrega de paquetes, permitiendo evaluar el rendimiento global de la red móvil.

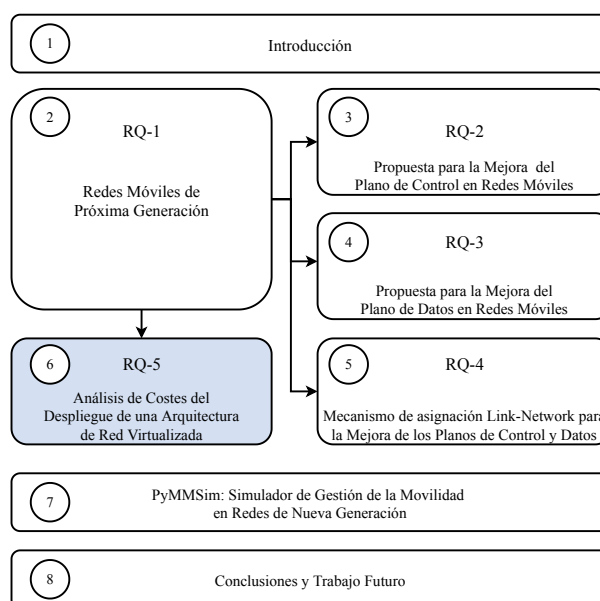
Los resultados obtenidos demuestran que nuestro algoritmo LNA puede reducir con éxito los costes de señalización hasta un 35 % en comparación con los algoritmos de referencia, sin penalizar el coste de entrega de paquetes que también se mejora, aunque en menor medida. Esta reducción de ambas métricas es lo que hace realmente interesante a la propuesta LNA, ya que las soluciones de gestión de la movilidad tienen como objetivo balancear la sobrecarga de señalización y los costes asociados al plano de datos, pero, tal y como se ha discutido en capítulos anteriores, generalmente la disminución de uno de los costes asociados a uno de los planos impacta negativamente en el otro, y viceversa. Además, otro aspecto que merece la pena ser destacado es que nuestra solución no requiere la modificación de ningún protocolo de gestión de movilidad involucrado en la comunicación.

Teniendo en cuenta las mejoras y beneficios introducidos por nuestra propuesta, los operadores de red móvil tienen la posibilidad de contemplar los aspectos mencionados anteriormente al planificar las arquitecturas de red de próxima generación, con el objetivo de mejorar el rendimiento de la red y diseñar redes para soportar los requisitos de los casos de uso emergentes en 5G. Por lo tanto, la asignación óptima entre estaciones base y la red de acceso se ha convertido en un desafío que deben abordar los operadores de redes móviles de nueva generación.

Capítulo 6

Análisis de Costes del Despliegue de una Arquitectura de Red Virtualizada

En este capítulo se lleva a cabo una evaluación de costes relacionados con el despliegue de tecnologías de red emergentes como SDN y NFV, analizando la viabilidad económica de dichos despliegues de red. Además, se propone un modelo analítico basado en costes de infraestructura (CAPEX) y costes de operación (OPEX), permitiendo obtener una evaluación comparativa y fiable entre los despliegues de red tradicionales y los despliegues de arquitecturas de red virtualizadas.



El análisis llevado a cabo ha tomado como base la Red Científico Tecnológica de Extremadura (RCT), de manera que se puedan cubrir sobre ella las necesidades de las redes de próxima generación, reduciendo costes y proporcionando agilidad en el despliegue de sus servicios.

Por tanto, el capítulo se organiza tal y como se presenta a continuación:

- La Sección 6.1 contextualiza el problema de la adaptación a la naturaleza dinámica de las redes de próxima generación por parte de los operadores de red, debido a la aparición de nuevos servicios requeridos por los usuarios móviles.
- La Sección 6.2 presenta la integración de la virtualización de red en el contexto de las redes de próxima generación.
- La Sección 6.3 describe el modelo de costes propuesto para analizar y evaluar de manera cuantitativa el despliegue de una arquitectura de red virtualizada basada en SDN/NFV.
- En la Sección 6.4 se muestran los resultados obtenidos, llevando a cabo una comparativa entre los costes necesarios para el despliegue de una arquitectura de red tradicional y una arquitectura de red basada en SDN/NFV, así como los costes asociados al despliegue del EPC (*Evolved Packet Core*) virtualizado.
- La Sección 6.5 muestra las conclusiones que han sido identificadas a partir del análisis realizado y la evaluación llevada a cabo.

6.1. Introducción

En los últimos años, el incremento exponencial del tráfico de datos móviles unido al despliegue de nuevos servicios sobre las redes actuales han propiciado que los operadores de red busquen nuevas tecnologías para adaptarse. Todo ello ha impulsado la creación de nuevos mecanismos y arquitecturas de gestión de redes eficientes, capaces de soportar tecnologías emergentes y nuevos servicios y aplicaciones móviles.

Por otro lado, todos estos cambios están teniendo un gran impacto en las estrategias económicas llevadas a cabo por los operadores de red. Cuando los operadores introducen nuevos servicios, expandiendo su infraestructura de red y/o optimizando sus recursos, tienen que tener en cuenta el coste de estas acciones. El aumento de los requisitos de eficiencia y disponibilidad por parte de los usuarios va ligado a un incremento de los costes de todo ello. Por lo tanto, las decisiones de planificación y diseño de las redes

deben tener en cuenta las estimaciones de costes con la mayor precisión posible. Para ello se utilizan modelos basados en CAPEX y OPEX [131, 132].

En definitiva, los operadores de red requieren de nuevos mecanismos y soluciones para cubrir de manera eficiente (técnicamente y económicamente) las necesidades que han ido apareciendo en estos entornos de red tan cambiantes. Estos mecanismos deben ser capaces de controlar y reservar dinámicamente los recursos de la red, para proporcionar la flexibilidad requerida por los operadores de red [4]. Por ello, tal y como se explicó en el Capítulo 2, las tecnologías SDN [27] y NFV [36] son vistas como una gran oportunidad para hacer frente a la naturaleza dinámica de las redes de próxima generación y al aprovisionamiento de servicios de manera flexible, proporcionando beneficios desde el punto de vista técnico, pero también económico; ya que estos enfoques tienen el potencial de conducir a importantes reducciones de costes de capital y costes de operación.

6.2. Virtualización de red

La tecnología de las Redes Definidas por Software surge como un nuevo paradigma de red, cuya principal característica es la separación del plano de datos del plano de control, con el objetivo de simplificar la gestión y configuración de la red [30]. La arquitectura propuesta por SDN es considerada como una importante oportunidad para gestionar las redes tradicionales, las cuales son complejas y difícilmente gestionables. SDN proporciona una vista global de la red a través de un controlador de red centralizado. Por lo tanto, el plano de control queda centralizado en el controlador de la red que, a su vez, gestiona el plano de datos a través de protocolos abiertos como OpenFlow [33]. SDN proporciona agilidad, permitiendo a los administradores de red una gestión dinámica de flujos y optimizando recursos ante las necesidades cambiantes de las aplicaciones, que pueden tener diferentes requerimientos en cuanto a características y calidad de servicio [133].

Por otro lado, la aparición del paradigma NFV ha sido muy importante desde el punto de vista del aprovisionamiento de servicios de telecomunicaciones. Este paradigma tiene como principal objetivo desacoplar las funciones de red de los dispositivos físicos en los cuales se ejecutan. Además, NFV tiene el potencial de proporcionar reducciones significativas de CAPEX y OPEX, y de facilitar y flexibilizar el despliegue de nuevos servicios con mayor agilidad [38], permitiendo alcanzar los requisitos de baja latencia y alta fiabilidad requerida por los servicios que se ofrecen en las redes 5G [39].

Tal y como se muestra en la Figura 6.1, los paradigmas SDN y NFV están íntimamente relacionados [37] y, con una integración eficiente de ambos paradigmas, se podría conseguir un importante ahorro de costes y una mayor flexibilidad en la provisión de servicios.

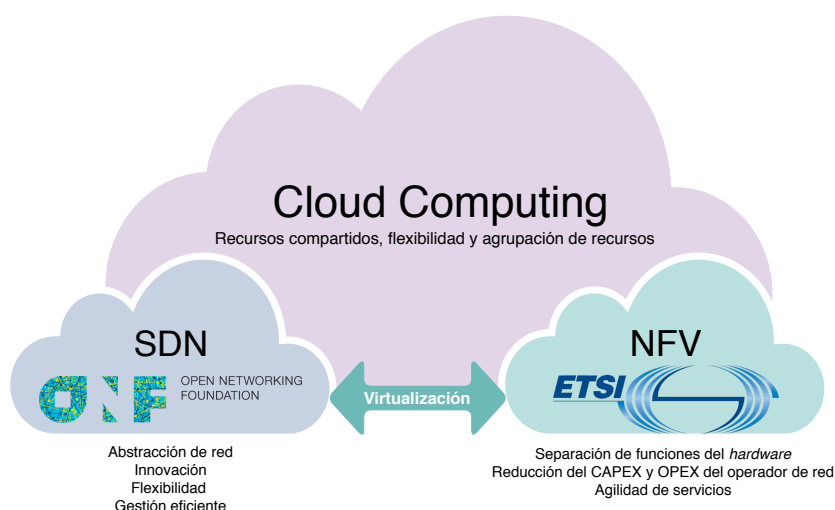


Figura 6.1: Relación entre SDN y NFV

La virtualización de red es un método en el que los recursos físicos de una red se dividen en segmentos o *slices*. Cada *slice* está aislada de las demás y pueden compartir infraestructura física en paralelo, con las ventajas que ello conlleva. Por ello, se acuñó el término de Cloud-RAN [134] que permite desacoplar el procesamiento de banda base de las propias estaciones base. Esta tecnología permite que el procesamiento se lleve a cabo en un centro de datos, reduciendo costes de forma significativa. Cloud-RAN ofrece un despliegue más fácil y flexible de nuevas tecnologías, a diferencia de las estaciones base (BS) tradicionales. De esta manera, aparece el concepto de SBS (BS definida por *software*), en la que un número determinado de BS virtuales pueden ser desplegadas simultáneamente.

Numerosos estudios establecen que SDN y NFV conducen a una reducción significativa del CAPEX para los operadores de red [132, 37]. En relación al OPEX [135], los procesos operativos automatizados podrían reducir la intervención humana, reduciendo los costes de personal y las operaciones de red que generen fallos. En lo que respecta al CAPEX, una provisión de funciones y servicios flexible, ágil y óptima puede reducir los costes de equipamiento y posponer las inversiones [136].

6.3. Modelo de costes

En esta sección se presenta el modelo de costes que se utilizará posteriormente, para analizar y evaluar de manera cuantitativa el despliegue de una arquitectura de red virtualizada.

El modelo utilizado establece que los costes totales de un determinado operador pueden ser divididos en dos, principalmente: CAPEX y OPEX [131, 132].

- Costes de capital (CAPEX). También llamados inversiones en bienes de capitales, ya que son inversiones de capital que crean beneficios. El CAPEX está íntimamente ligado a los costes de la infraestructura fija de la empresa y se amortizan con el tiempo. En general, el CAPEX empieza a tener sentido cuando una empresa invierte en una determinada compra de un activo fijo o para añadir valor a un activo existente con una vida útil que se extiende más allá del año actual. Para un operador de red, el CAPEX incluye los costes relacionados con la compra de terrenos y/o edificios (por ejemplo, para alojar el equipamiento tecnológico), la infraestructura de red y el despliegue inicial de dicha infraestructura. Hay que tener en cuenta que la compra de equipamiento siempre tiene que ser incluida en el CAPEX, independientemente de si el pago se realiza de una vez o si se extiende a lo largo del tiempo.
- Costes de operación (OPEX). Se trata de un coste permanente para el funcionamiento de un producto, negocio o sistema. También llamados costes de funcionamiento o costes operativos. El OPEX no contribuye a los costes del despliegue de la infraestructura; representan el coste necesario para mantener operativa la infraestructura, incluyendo los costes de operaciones técnicas y comerciales, administración, etc. Para un operador de red, el OPEX está formado por los costes de alquiler (edificio, terreno, equipos de red, etc), los costes de energía y los costes de mantenimiento, administración y operación, entre otros.

Las estimaciones del CAPEX y OPEX realizadas se han llevado a cabo utilizando el modelo que se presenta en la Figura 6.2. En ella se pueden observar los costes que han sido considerados durante el estudio y análisis.

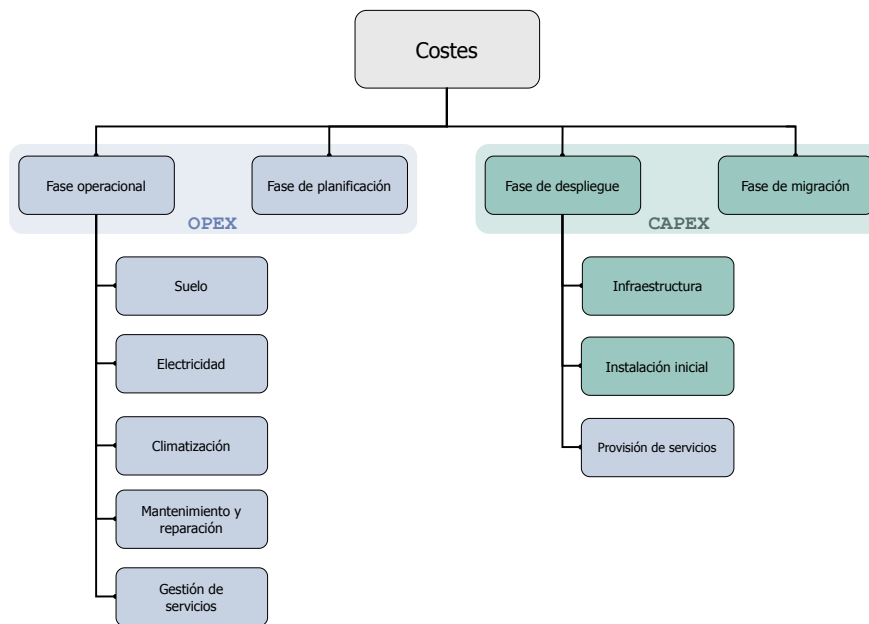


Figura 6.2: Costes de infraestructura y operación considerados

Además, para poder evaluar y comparar convenientemente los costes del despliegue de una arquitectura de red virtualizada, que haga uso de los nuevos paradigmas emergentes, es necesario analizar también los costes a los que los operadores tendrían que hacer frente si se usara una arquitectura de red tradicional. Desde el punto de vista de la tecnología, la Figura 6.3 muestra una comparativa entre ellas.

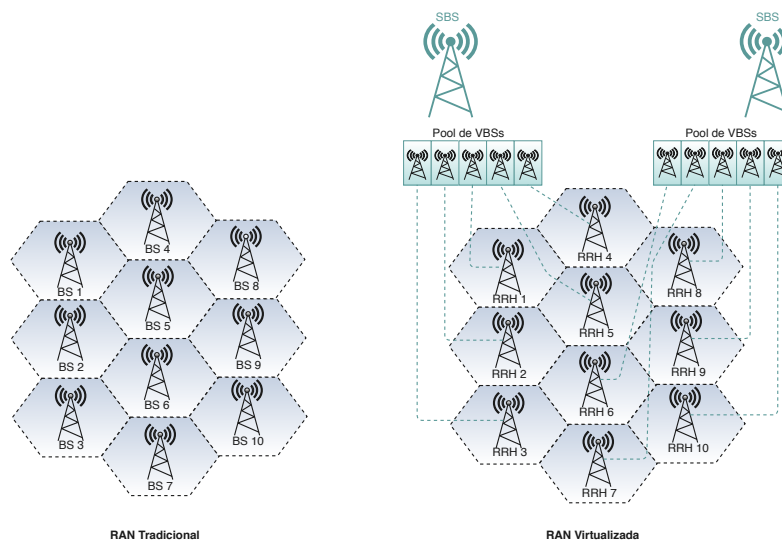


Figura 6.3: Comparativa entre una RAN tradicional y una RAN virtualizada (C-RAN)

En realidad, una RAN (*Radio Access Networks*) virtualizada (C-RAN) se trata de una evolución de una RAN tradicional, en la que el elemento principal pasa a ser una BS virtualizada (VBS), es decir, un número determinado de BS virtuales pueden ser desplegadas simultáneamente sobre una BS física [137]. La estación base que albergue varias VBS la denominaremos SBS. Concretamente, una C-RAN consta de tres componentes principales: la antena, que junto al RRH (*Remote Radio Head*) se ubican en una localización remota y estarán controlados por las VBS; el propio *pool* de VBS, que está compuesto por procesadores de alto rendimiento que hacen uso de tecnologías de virtualización en tiempo real y una red de baja latencia para conectar las antenas al *pool* de VBS.

6.3.1. CAPEX y OPEX en una red de acceso tradicional

En esta sección se modelan los costes relacionados con una red de acceso radio tradicional (RAN), que nos permitirán evaluar el impacto relativo al despliegue de la infraestructura de red necesaria y su mantenimiento como si fuera un operador de red tradicional.

Se asume que en un área A existe un determinado número de operadores (N_{op}) que se encuentran dando cobertura a un conjunto de clientes. El número de estaciones base requeridas para cubrir el área A depende, por un lado, del radio de cobertura de la estación base (R_{max}) y, por otro lado, del número total de usuarios en el área A (N_{UT}). Por lo tanto, el número total de estaciones base requeridas por cada operador viene dado por la Ecuación 6.1.

$$N_{BSO} = \frac{A}{\pi \cdot R_{max}^2} \quad (6.1)$$

Si λ refleja la densidad de usuarios por unidad de área, se puede definir el número total de usuarios por operador (N_{UOp}) como se muestra en la Ecuación 6.2.

$$N_{UOp} = \lambda \cdot A = \pi \cdot \lambda \cdot N_{BSO} \cdot R_{max}^2 \quad (6.2)$$

Siendo C_{dBS} los costes asociados al despliegue inicial de una estación base y C_{BS} los costes de cada estación base (equipamiento), se puede definir el CAPEX total de una arquitectura RAN tradicional como se muestra en la Ecuación 6.3.

$$CAPEX(RAN) = \sum_{i=1}^{N_{op}} \sum_{j=1}^{N_{BSO_i}} \left(C_{dBS} + C_{BS} \right)_{ij} \quad (6.3)$$

Obteniendo N_{BSO} de la Ecuación 6.2 y, aplicándolo a la Ecuación 6.3, se obtiene el coste de la infraestructura para todos los usuarios (C_{infra}), tal y como se muestra en la Ecuación 6.4.

$$C_{infra} = \overbrace{N_{op} \cdot N_{UOp}}^{N_{UT}} \cdot \frac{C_{dBS} + C_{BS}}{\pi \cdot \lambda \cdot R_{max}^2} \quad (6.4)$$

Además, si se quisiera conocer el coste de la infraestructura por usuario ($C_{infra-u}$), dividiremos la Ecuación 6.4 entre el número de usuarios totales en la red (N_{UT}), tal y como muestra la Ecuación 6.5.

$$C_{infra-u} = \frac{C_{infra}}{N_{UT}} = \frac{C_{dBS} + C_{BS}}{\pi \cdot \lambda \cdot R_{max}^2} \quad (6.5)$$

Por otro lado, la Ecuación 6.6 define el OPEX total de una arquitectura RAN tradicional. Para ello se tendrán en cuenta la energía consumida por la BS (P_{BS}), el coste del kWh (C_{kWh}), los costes asociados a la operación, mantenimiento y administración ($C_{OA\&M}$) y los costes de alquiler (C_{ALQ}).

$$OPEX(RAN) = \sum_{i=1}^{N_{op}} \sum_{j=1}^{N_{BSO_i}} \left(P_{BS} \cdot C_{kWh} + C_{OA\&M} + C_{ALQ} \right)_{ij} \quad (6.6)$$

P_{BS} engloba la potencia de todas las antenas que conforman la estación base. Considerando que una estación base está formada por un número de antenas determinado (N_a), con una potencia determinada cada una (P_a), se puede estimar la potencia eléctrica total de cada BS, tal y como se muestra en la Ecuación 6.7.

$$P_{BS} = N_a \cdot P_a \quad (6.7)$$

Por último, se define $TCO(RAN)$ como el coste total del operador, que resulta de sumar los dos costes definidos anteriormente, tal y como se muestra en la Ecuación 6.8.

$$\begin{aligned}
TCO(RAN) &= CAPEX(RAN) + OPEX(RAN) = \\
&= \sum_{i=1}^{N_{op}} \sum_{j=1}^{N_{BSO_i}} \left(C_{dBS} + C_{BS} \right)_{ij} + \sum_{i=1}^{N_{op}} \sum_{j=1}^{N_{BSO_i}} \left(P_{BS} \cdot C_{kWh} + C_{OA\&M} + C_{ALQ} \right)_{ij} \quad (6.8)
\end{aligned}$$

6.3.2. CAPEX y OPEX en una red de acceso virtualizada (C-RAN)

A continuación se modelan los costes relacionados con una red de acceso radio virtualizada, que permitirán evaluar los costes relativos a la infraestructura de red necesaria para desplegar un operador de red virtualizado.

Se asume que en un área A existe un determinado número de estaciones base virtuales VBS, desplegadas sobre la SBS, que se encuentran dando cobertura a un conjunto de usuarios. El número de SBS requerido para cubrir el área A estará determinado por el radio de cobertura de la SBS (R_{max}) y por el número total de usuarios en el área A (N_{UT}), tal y como se muestra en la Ecuación 6.9.

$$N_{SBS} = \frac{A}{\pi \cdot R_{max}^2} \quad (6.9)$$

Si λ refleja la densidad de usuarios por unidad de área y N_{sl} define el número de *slices* (VBS) por cada SBS, se tiene que el número total de usuarios en un área A es N_{UT} (Ecuación 6.10).

$$N_{UT} = \lambda \cdot A = \sum_{i=1}^{N_{sl}} \left(\pi \cdot \lambda \cdot N_{SBS} \cdot R_{max}^2 \right) \quad (6.10)$$

Además, C_{SBS} representa el coste de una SBS. Se asume que el coste total de una SBS aumenta linealmente con respecto al número de VBS desplegadas en ella, tal y como se muestra en la Ecuación 6.11 [138].

$$C_{SBS} = C_{BS} \cdot (1 + 0,2 \cdot (N_{sl} - 1)) \quad (6.11)$$

Siendo C_{dSBS} el coste asociado al despliegue inicial de una SBS, se puede definir el CAPEX total de una arquitectura C-RAN como se muestra en la Ecuación 6.12.

$$CAPEX(C - RAN) = C_{dSBS-total} + C_{SBS-total} = N_{SBS} \cdot (C_{dSBS} + C_{SBS}) \quad (6.12)$$

Obteniendo N_{SBS} de la Ecuación 6.10 y, aplicándolo a la Ecuación 6.12, se obtiene el coste de la infraestructura para todos los usuarios ($C_{infra-sbs}$), tal y como se muestra en la Ecuación 6.13.

$$C_{infra-sbs} = \frac{N_{UT}}{N_{sl} \cdot \pi \cdot \lambda \cdot R_{max}^2} \cdot (C_{dSBS} + C_{SBS}) \quad (6.13)$$

Además, el coste de la infraestructura por usuario ($C_{infra-sbs-u}$) queda definido por la Ecuación 6.14.

$$C_{infra-sbs-u} = \frac{C_{dSBS} + C_{SBS}}{N_{sl} \cdot \pi \cdot \lambda \cdot R_{max}^2} \quad (6.14)$$

Por otro lado, la Ecuación 6.15 define el OPEX para una arquitectura C-RAN. Estará formado por los costes derivados del consumo energético de las SBS, que serán calculados a partir de la potencia de la SBS (P_{SBS}) y el coste del kWh (C_{kWh}). Además, también se tendrán en cuenta costes asociados a la operación, mantenimiento y administración ($C_{OA\&M}$) y los costes de alquiler (C_{ALQ}). Hay que tener en cuenta que cada SBS tiene más capacidades que una BS tradicional, por lo que P_{SBS} dependerá del número de VBS que se desplieguen sobre ella.

$$OPEX(C - RAN) = \sum_{i=1}^{N_{SBS}} \left(P_{SBS} \cdot C_{kWh} + C_{OA\&M} + C_{ALQ} \right)_i \quad (6.15)$$

Considerando que cada VBS está formada por un número de antenas determinado (N_a), se puede estimar la potencia eléctrica de cada SBS (P_{SBS}) como se indica en la Ecuación 6.16.

$$P_{SBS} = N_{sl} \cdot N_a \cdot P_a \quad (6.16)$$

Por último, se define el $TCO(C - RAN)$ como el coste total del operador relativo al despliegue de una RAN virtualizada. En la Ecuación 6.17 se define este parámetro.

$$\begin{aligned}
TCO(C - RAN) &= CAPEX(C - RAN) + OPEX(C - RAN) = \\
&= \sum_{i=1}^{N_{SBS}} \left(C_{dSBS} + C_{SBS} \right) + \sum_{i=1}^{N_{SBS}} \left(P_{SBS} \cdot C_{kWh} + C_{OA\&M} + C_{ALQ} \right)_i \quad (6.17)
\end{aligned}$$

6.3.3. CAPEX y OPEX en una red de acceso 3GPP virtualizada

Todavía se puede profundizar un poco más y aprovechar aún más los beneficios que proporcionan la tecnología NFV, virtualizando los elementos principales del núcleo de una red de acceso 3GPP (EPC), que está formado por las siguientes entidades:

- MME, *Mobility Management Entity*. Es el nodo de control clave para el acceso a la red. Es responsable del seguimiento de usuario móvil. También se encarga de la gestión de la movilidad y de las sesiones cuando se produce un *handover* o traspaso entre celdas de cobertura, y se encarga de la autenticación del usuario interactuando con el HSS.
- HSS, *Home Subscriber Server*. Se trata de una base de datos central con la capacidad de almacenar información de los usuarios, localización y datos del servicio.
- SGW, *Serving Gateway*. Esta entidad se encarga del intercambio de tráfico de usuario entre la red de acceso y el núcleo de red IP. Cuando un usuario móvil se mueve a través de diferentes antenas, el SGW sirve como ancla de movilidad, interceptando los paquetes que son dirigidos al nodo móvil para posteriormente ser enviados a la ubicación actual del mismo.
- PGW, *Packet Data Network Gateway*. Esta entidad es la encargada del intercambio de tráfico con redes externas (PDN, *Packet Data Network*). Funciona como punto de anclaje para la salida a PDN externas y proporciona características de filtrado de paquetes y recogida de datos de tarificación, entre otras.

Siendo N_{SERV} el número total de servidores necesarios para poder realizar la virtualización de red y C_{SERV} el coste de un servidor, se puede establecer que el coste total de adquisición de todos los servidores que son necesarios para operar convenientemente viene dado por la Ecuación 6.18.

$$C_{SERV-total} = N_{SERV} \cdot C_{SERV} \quad (6.18)$$

Por lo tanto, se puede definir el CAPEX total para virtualizar el EPC como $CAPEX(vEPC)$, tal y como se muestra en la Ecuación 6.19, teniendo en cuenta el coste total de los servidores requeridos para la virtualización ($C_{SERV-total}$), los costes de licencia *software* (C_{LIC}) y los costes asociados al despliegue e instalación inicial (C_{INI}).

$$CAPEX(vEPC) = C_{SERV-total} + C_{LIC} + C_{INI} \quad (6.19)$$

Por otro lado, para definir el OPEX del EPC virtualizado ($OPEX(vEPC)$) se tendrá en cuenta el consumo energético de todos los servidores necesarios para alojar el número requerido de máquinas virtuales. Para virtualizar cualquier entidad del EPC (E_{EPC}), el número de máquinas virtuales necesarias para dicha demanda se determinará mediante el requisito máximo de CPU, memoria, red o almacenamiento, tal y como se muestra en la Ecuación 6.20.

$$VMS_{E_{EPC}} = \max \left(VM_{CPU-E_{EPC}}, VM_{RAM-E_{EPC}}, VM_{ALM-E_{EPC}}, VM_{RED-E_{EPC}} \right) \quad (6.20)$$

Los requisitos de CPU, memoria RAM, almacenamiento y red son definidos tal y como se describe en las Ecuaciones 6.21, 6.22, 6.23 y 6.24, respectivamente.

$$VM_{CPU-E_{EPC}} = \left\lceil \frac{CPU_{E_{EPC}}}{CPU_{porVM}} \right\rceil \quad (6.21)$$

$$VM_{RAM-E_{EPC}} = \left\lceil \frac{RAM_{E_{EPC}}}{RAM_{porVM}} \right\rceil \quad (6.22)$$

$$VM_{ALM-E_{EPC}} = \left\lceil \frac{ALM_{E_{EPC}}}{ALM_{porVM}} \right\rceil \quad (6.23)$$

$$VM_{RED-E_{EPC}} = \left\lceil \frac{RED_{E_{EPC}}}{RED_{porVM}} \right\rceil \quad (6.24)$$

Siendo $VMS_{porSERV}$ el número de máquinas virtuales que puede contener un servidor físico, el número de servidores requeridos para albergar cada entidad del EPC viene dado por la Ecuación 6.25.

$$N_{SERV-EPC} = \frac{VMS_{EPC}}{VMS_{porSERV}} \quad (6.25)$$

Finalmente, la Ecuación 6.26 define el OPEX total para el EPC virtual, siendo P_{SERV} la potencia de un servidor y C_{kWh} el coste del kWh.

$$OPEX(vEPC) = \sum_{i \in E_{EPC}} \left(N_{SERV_i} \right) \cdot P_{SERV} \cdot C_{kWh} \quad (6.26)$$

6.4. Caso RCT: resultados obtenidos

Extremadura cuenta con una infraestructura de fibra óptica a la que la Junta de Extremadura denominó como Red Científico Tecnológica de Extremadura (RCT). Tiene como principal objetivo proporcionar servicios de comunicaciones avanzadas y de altas prestaciones a los centros tecnológicos y de investigación ubicados en la región.

Todo ello surgió ante la necesidad de la Comunidad Autónoma de Extremadura de contar con una infraestructura de telecomunicaciones tecnológicamente avanzada que soportara accesos telemáticos de última generación. Por ello, se acometieron distintos proyectos cuya finalidad fue el tendido de una red de fibra óptica que permitiera conformar una red troncal de telecomunicaciones de alta velocidad que cumpliera con los siguientes objetivos:

- Unir las infraestructuras de los campus y/o edificios de la Universidad de Extremadura.
- Interconectar los centros sanitarios de referencia, tecnológicos y principales centros administrativos de la región.
- Dar cobertura a proyectos e iniciativas impulsadas por la Administración Regional en el ámbito de la investigación científica e innovación tecnológica.
- Fomentar el intercambio de información y conocimiento entre las universidades, centros de investigación y centros tecnológicos, llegando a ser la columna vertebral del desarrollo tecnológico en la región.

- Permitir la conexión con otras redes de investigación de ámbito nacional y europeo.

Dada la gran capacidad de transmisión de datos que ofrece la Red Científico Tecnológica de Extremadura, se exceden las necesidades de los centros tecnológicos conectados inicialmente, bajo cuya filosofía surgió la necesidad del despliegue de dicha red. Por ello, se puede reutilizar el sobredimensionamiento de red para la prestación de servicios de acceso a las infraestructuras de telecomunicaciones tanto a clientes públicos como privados. Con esta infraestructura, se dispondría de una red troncal de telecomunicaciones que conectaría los municipios de Extremadura y que podría ser utilizada como red troncal por un operador de telecomunicaciones para ofrecer servicios de banda ancha en dichos municipios, tal y como se muestra en la Figura 6.4.

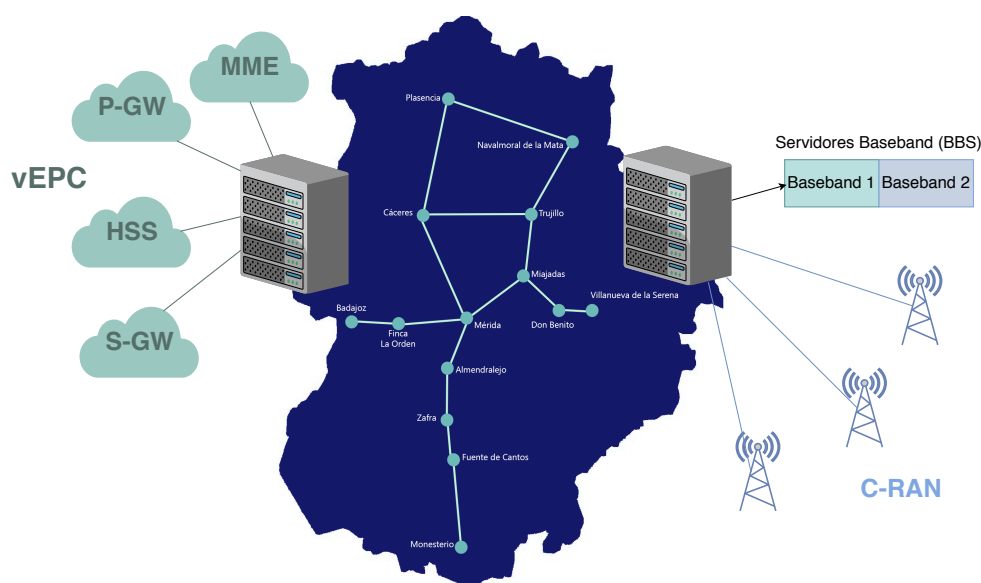


Figura 6.4: Red Científico Tecnológica virtualizada

Por lo tanto, en esta sección se evaluará de forma cuantitativa la implantación de mecanismos de nueva generación en la RCT, con el objetivo de cuantificar los costes operacionales y de capital.

La Tabla 6.1 muestra los parámetros por defecto utilizados para la evaluación del CAPEX, OPEX y TCO, así como para el despliegue del EPC virtualizado. La elección de estos parámetros se ha realizado en base a estudios realizados anteriormente [136, 139, 140, 141].

Tabla 6.1: *Parámetros tenidos en cuenta para la evaluación*

Parámetro	Definición	Valor
C_{dBS}	Costes asociados al despliegue inicial de una BS	5 000 €
C_{dSBS}	Costes asociados al despliegue inicial de una SBS	5 000 €
C_{BS}	Coste de una BS (equipamiento)	15 596 €
P_a	Potencia de cada antena en cada BS	615 W
N_a	Número de antenas en una BS	4
C_{kWh}	Coste del kWh	0,14 €
$C_{OA\&M}$	Costes de operación, mantenimiento y administración	4 000 €/año
C_{ALQ}	Coste de alquiler para el emplazamiento de cada BS	1 000 €/año
R_{max}	Radio de cobertura de una BS	2 km
C_{LIC}	Costes de licencias <i>software</i>	5 000 €
C_{INI}	Costes de instalación inicial	30 000 €
C_{SERV}	Coste de un servidor	5 500 €
P_{SERV}	Potencia de un servidor	400 W
CPU_{EPC}	CPU requerida para virtualizar una determinada entidad del EPC	6 cores
RAM_{EPC}	Memoria RAM requerida para virtualizar una determinada entidad del EPC	40 GB
ALM_{EPC}	Almacenamiento en disco requerido para virtualizar una determinada entidad del EPC	1 TB
RED_{EPC}	Especificaciones de red requeridas para virtualizar una determinada entidad del EPC	30 Gpbs
CPU_{porVM}	CPU por cada máquina virtual	2 cores
RAM_{porVM}	Memoria RAM por cada máquina virtual	8 GB
ALM_{porVM}	Almacenamiento por cada máquina virtual	150 GB
RED_{porVM}	Red por máquina virtual	40 Gpbs
$VMS_{porSERV}$	Número de máquinas virtuales que se pueden desplegar sobre un servidor físico	10

Además, los valores de referencia para el número de estaciones base han sido: 10, 20, 30, 50 y 100. Para ello se ha tenido en cuenta la información geográfica de las estaciones base en la Comunidad Autónoma de Extremadura [142]. En el caso del escenario C-RAN se ha considerado que sobre cada SBS se pueden desplegar hasta 5 VBS.

En primer lugar se han calculado los costes de capital en relación al número de estaciones base desplegadas, teniendo en cuenta los valores establecidos en la Tabla 6.1. La Figura 6.5 muestra la comparativa de los costes de capital entre los dos tipos de arquitectura desplegadas (RAN y C-RAN). En ambos casos, el CAPEX aumenta de manera lineal con respecto al número de estaciones base desplegadas. Tal y como se muestra en la Figura 6.5, se pueden conseguir ahorros de los costes de capital del orden del 70 % cuando se despliegan 100 BS bajo una arquitectura C-RAN, con respecto a una arquitectura RAN tradicional.

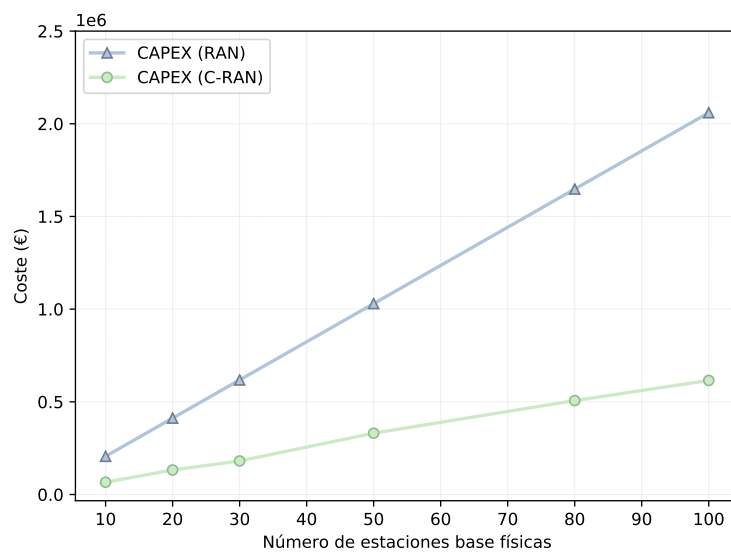


Figura 6.5: Análisis de CAPEX: RAN vs C-RAN

Por otro lado, la Figura 6.6 refleja la evolución de los costes de operación con respecto al número de estaciones base físicas desplegadas en un área determinada. En el caso del OPEX, al ser costes relacionados con el mantenimiento de la infraestructura, en este estudio se han considerado los costes asociados al primer año de operación, es decir, OPEX anual. En el caso de C-RAN se consiguen ahorros de OPEX de aproximadamente un 60 % con respecto a la arquitectura RAN.

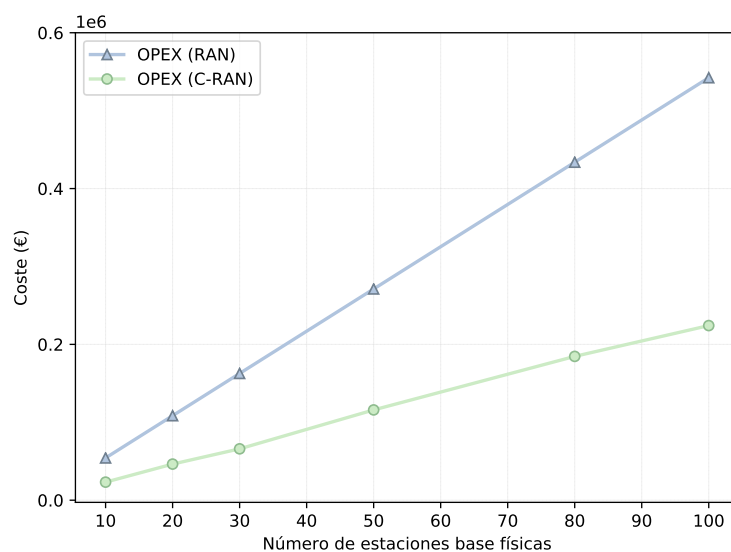


Figura 6.6: Análisis de OPEX: RAN vs C-RAN

La Figura 6.7 muestra la influencia del número de estaciones base desplegadas en el coste total del operador de red. Además, las reducciones de CAPEX y OPEX en el caso de la arquitectura C-RAN se ven reflejadas en el TCO. Este coste total está muy influenciado por los costes de capital para los dos despliegues analizados.

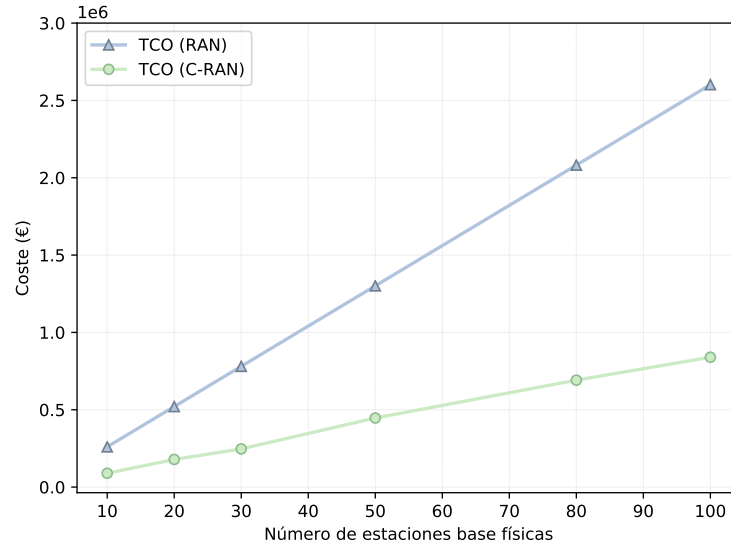


Figura 6.7: Análisis de TCO: RAN vs C-RAN

Además, se ha calculado el coste de la infraestructura en función de la densidad de usuarios en una determinada área, tal y como se muestra en la Figura 6.8.

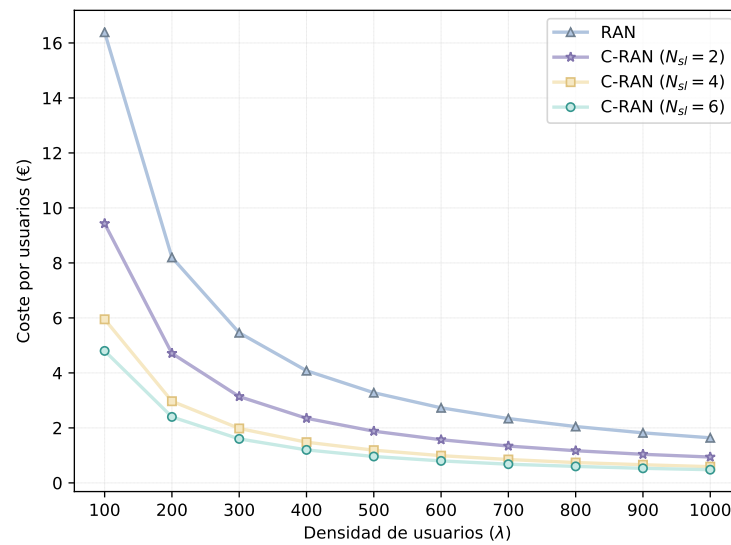


Figura 6.8: Comparativa de costes relativos de RAN vs C-RAN en función de la densidad de usuarios

Para el caso de la arquitectura C-RAN, a medida que el número de *slices* por SBS aumenta, el coste de la inversión en infraestructura por usuario disminuye. Además, el coste por usuarios en la RAN tradicional es mucho mayor que el coste por usuarios en la arquitectura C-RAN. Cuando hay una mayor densidad de usuarios en un área determinada, los costes se disminuyen considerablemente, pero siempre los del despliegue de red tradicional están por encima de los del despliegue de red virtualizado.

Finalmente, se ha evaluado el despliegue del EPC virtualizado, calculando los costes de capital y de operación. Para ello se han aplicado las Ecuaciones 6.19 y 6.26, tomando como valores de referencia los mostrados en la Tabla 6.1. La Figura 6.9 muestra la evolución de los costes, que aumentan de manera lineal con respecto al número de EPC virtuales desplegados (con sus cuatro entidades).

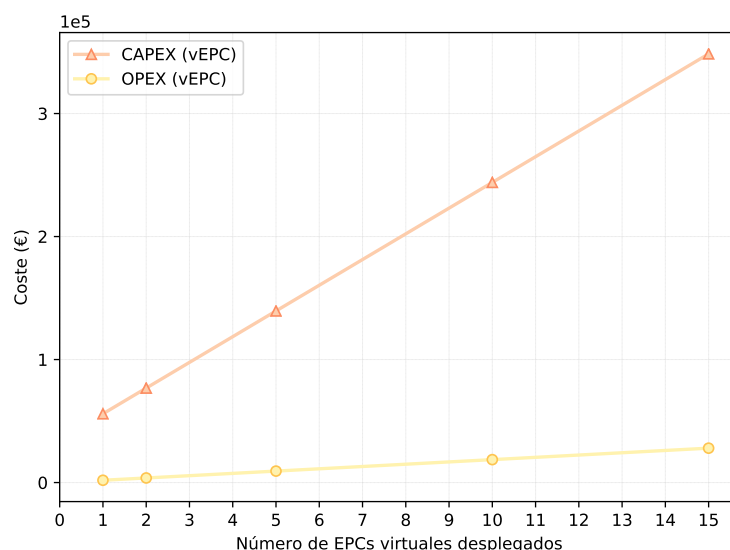


Figura 6.9: Análisis del CAPEX y OPEX del EPC virtualizado

6.5. Conclusiones

Debido al rápido crecimiento del tráfico de datos móviles y los nuevos servicios demandados por los usuarios, los proveedores de servicios de telecomunicaciones han propuesto nuevas formas de prestación de servicios en términos de flexibilidad, agilidad y ahorro de costes a través de la gestión y ahorro de los costes de operación y los costes de capital.

Tecnologías como NFV y SDN han surgido para permitir que la infraestructura de red sea más flexible y abierta. Además, los proveedores de servicios pueden

implementar servicios innovadores que reduzcan los costes de operación, administración y mantenimiento.

SDN y NFV están estrechamente relacionados y todas estas tecnologías van a ser muy importantes en los entornos de prestación de servicios, dado que estas arquitecturas están definidas e integradas con redes de próxima generación o redes 5G.

En este capítulo se propone un enfoque para reducir el impacto de NFV y SDN en el coste total de despliegue y mantenimiento para un operador de red. Se presenta un modelo de análisis de costes que proporciona una visión técnica y económica para decidir qué arquitectura es la más acertada a la hora de desplegar las tecnologías de red y los servicios que han de ser desarrollados. Además, se establece un análisis cuantitativo para evaluar los beneficios y limitaciones de estas tecnologías emergentes.

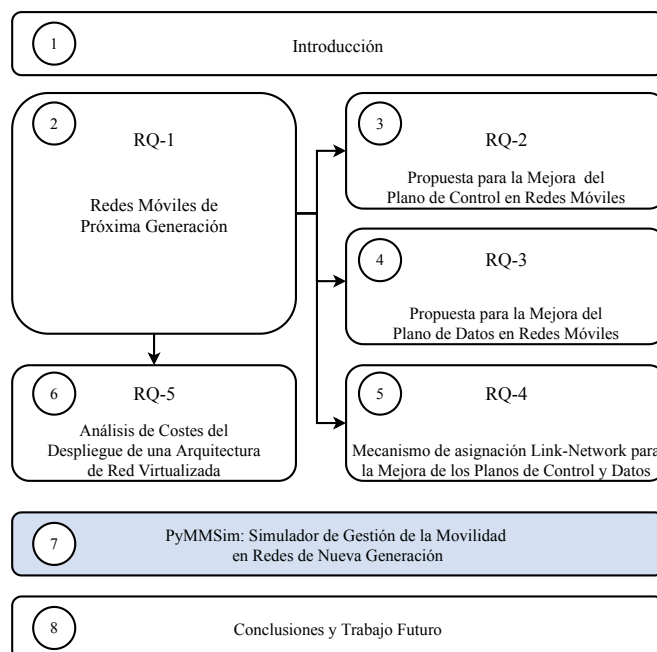
El modelo cuantitativo se utiliza para realizar experimentos que muestran una reducción significativa de los costes cuando se aplican mecanismos de virtualización en las redes de acceso. El análisis realizado revela que la arquitectura propuesta proporciona importantes ahorros de CAPEX, OPEX, TCO. Los resultados experimentales obtenidos de la comparativa llevada a cabo muestran que:

- El CAPEX para la red de acceso virtualizada se puede reducir hasta un 70 por ciento con respecto al caso tradicional.
- En comparación con la red de acceso tradicional, el OPEX para la red de acceso virtualizada puede reducirse hasta un 59 por ciento.
- El TCO para la arquitectura virtualizada se puede reducir hasta un 68 por ciento.
- La reducción de costes aumenta cuando mayor es el número de particiones virtuales.
- Además esta reducción es mayor cuando el EPC es virtualizado.

Capítulo 7

PyMMSim: Simulador de Gestión de la Movilidad en Redes de Nueva Generación

En este capítulo se describen los aspectos claves del diseño y la implementación de un simulador de gestión de la movilidad en redes de nueva generación (PyMMSim) desarrollado en Python, que ha sido utilizado para obtener los resultados de simulación presentados en esta Tesis.



Con la implementación de este simulador hemos conseguido simplificar la generación de resultados de movilidad en gran medida, permitiendo evaluar la escalabilidad de los diferentes protocolos de movilidad y las propuestas llevadas a cabo en esta Tesis. Además, se ha propuesto un diseño modular que permite realizar ampliaciones del mismo, añadiendo nuevas funcionalidades fácilmente, como por ejemplo nuevos modelos de movimiento, nuevos modelos de tráfico o nuevos protocolos de gestión de la movilidad, entre otros.

El capítulo se organiza tal y como se presenta a continuación:

- La Sección 7.1 describe la importancia de la simulación en entornos complejos de red, además de introducir los aspectos claves del simulador desarrollado.
- La Sección 7.2 presenta la estructura general y las características principales de PyMMSim, describiendo en detalle los módulos más importantes que lo componen.
- En la Sección 7.3 se presentan las características principales del escenario de simulación genérico utilizado en la mayor parte de las evaluaciones.
- La Sección 7.4 describe el entorno de ejecución sobre el cual ha sido desplegado PyMMSim, permitiendo evaluar la escalabilidad de las propuestas desarrolladas.
- En la Sección 7.5 se definen las métricas que se han contemplado e introducido en el simulador para evaluar la gestión de la movilidad.
- La Sección 7.6 muestra las conclusiones que han sido identificadas tras llevar a cabo el diseño y la implementación de nuestro simulador.

7.1. Introducción

Una fase crucial durante el diseño e implementación de sistemas de comunicaciones (protocolos, algoritmos o arquitecturas de red), es la evaluación del rendimiento y la comprensión del comportamiento de los sistemas y sus componentes. En general, para esta fase existen tres técnicas bien diferenciadas entre sí: modelo analítico, simulación o evaluación experimental. Cada una de estas técnicas cuenta con una serie de ventajas e inconvenientes y, además, en la literatura hay muchos estudios que discuten sobre cuándo usar una técnica u otra, su aplicación y las dificultades relacionadas con cada una de ellas [143, 144, 145]. Law y Kelton [143] definen las formas de estudiar un sistema tal y como se muestra en la Figura 7.1.

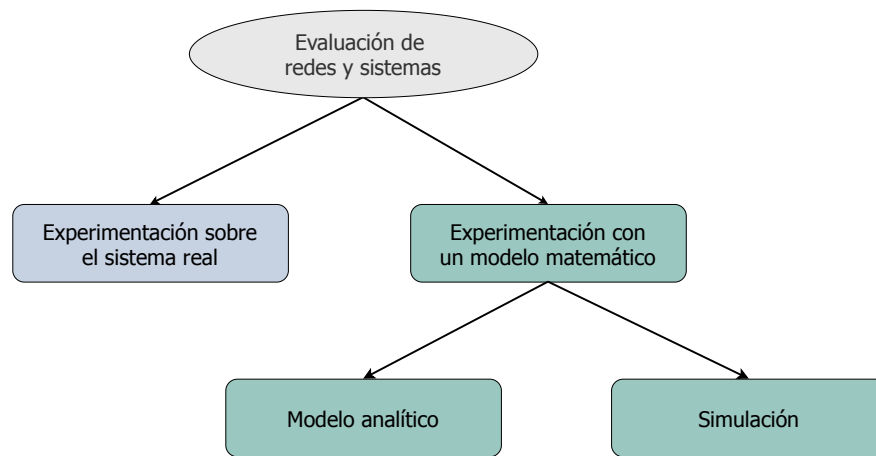


Figura 7.1: Maneras de estudiar un sistema

Por lo tanto, una cuestión importante a la hora de evaluar el rendimiento de un sistema de comunicaciones es la decisión de llevar a cabo la evaluación a través de experimentaciones sobre el sistema real o a través de modelos matemáticos (analíticos o simulación). La evaluación experimental requiere de la implementación de un prototipo que implica unos costes y un esfuerzo generalmente altos. Por ello, a menudo se aplican modelos analíticos o simulados, los cuales se basan en un modelo que representa el sistema de la forma más precisa posible. Sin embargo, en ocasiones, la simulación puede ser la única alternativa cuando se pretende estudiar un sistema complejo que está formado por elementos estocásticos (no deterministas), que no pueden ser tratados con la precisión necesaria en un modelo analítico, ya que estos modelos matemáticos detallados se vuelven intratables.

En definitiva, la simulación permite evaluar el rendimiento de un sistema bajo diferentes condiciones de operación con relativa facilidad, *imitando* el comportamiento de un sistema del mundo real [146, 147]. En otras ocasiones, se aplica para comparar diferentes alternativas de diseño que tengan que satisfacer unos requerimientos específicos, o incluso para optimizar un determinado diseño. Estas simulaciones son aplicadas en diferentes campos de investigación y existen diferentes tipos. En concreto, en el campo de las redes de comunicaciones la técnica de simulación más utilizada es la simulación de eventos discretos, cuya principal característica es que el estado del modelo de simulación es actualizado solamente cuando ocurre algún evento; mientras tanto el sistema permanece igual [146]. La razón detrás del éxito de este tipo de simulación en el contexto de las redes de comunicaciones, es que el paradigma se adapta muy bien a los sistemas considerados, proporcionando una manera simple y flexible de evaluar los enfoques de red y estudiar su comportamiento bajo diferentes condiciones [144]. La Figura 7.2 muestra la evolución en el tiempo de una simulación de eventos discretos.

Tal y como se puede observar, cada evento e_i ocurre en un tiempo determinado t_i , produciéndose un cambio en el estado del sistema.

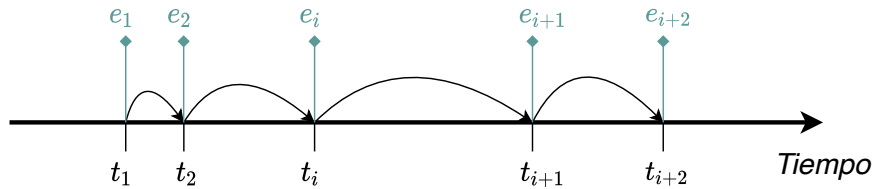


Figura 7.2: Principio de la simulación de eventos discretos. Durante la simulación el estado del sistema cambia solo en puntos discretos de tiempo t_i .

Teniendo en cuenta los principios de un simulador de eventos discretos, se puede definir el algoritmo genérico de tratamiento de eventos tal y como se muestra en la Figura 7.3, el cual está formado por tres fases principalmente: inicialización, procesamiento de eventos y generación de la salida.

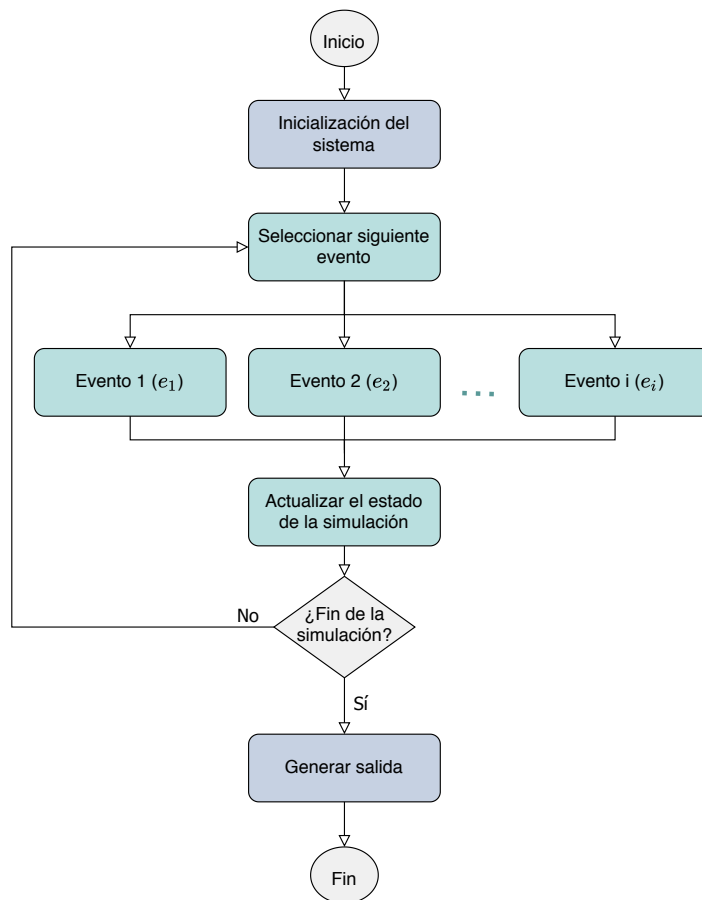


Figura 7.3: Diagrama de flujo del algoritmo genérico de tratamiento de eventos discretos

Durante la simulación, el estado del sistema evoluciona con el tiempo hasta que se alcanza el tiempo total de simulación y, por lo tanto, todos los eventos habrán sido tratados. La lista que contiene todos los eventos se ordena de menor a mayor de acuerdo al tiempo de ocurrencia: $listaEventos = \{e_1[t_1], e_2[t_2], \dots, e_i[t_i]\}$ donde $t_1 \leq t_2 \leq \dots \leq t_i$.

7.2. Estructura de PyMMSim

Una vez que se han descrito las características principales y el modo de funcionamiento general de un simulador de eventos discretos, en esta sección se definirá la estructura y los detalles de nuestro simulador (PyMMSim).

El simulador de eventos discretos PyMMSim se ha desarrollado con el principal objetivo de simular y analizar el proceso de gestión de la movilidad en redes de nueva generación. La idea que subyace bajo PyMMSim es la propuesta de un diseño modular que permita evaluar el proceso de movilidad, facilitando el desarrollo de nuevas funcionalidades. La Figura 7.4 muestra el diseño modular del simulador PyMMSim. Tal y como se puede observar, está dividido en un núcleo central que se encarga de controlar la simulación en todo momento, además de interactuar con todos los módulos que lo componen. Actualmente, PyMMSim se sustenta sobre seis módulos principalmente que implementan el modelo de red, la geometría de red celular, el modelo de tráfico, los modelos de movimiento, los protocolos de gestión de la movilidad y los algoritmos de asociación entre el nivel físico y el nivel de red.

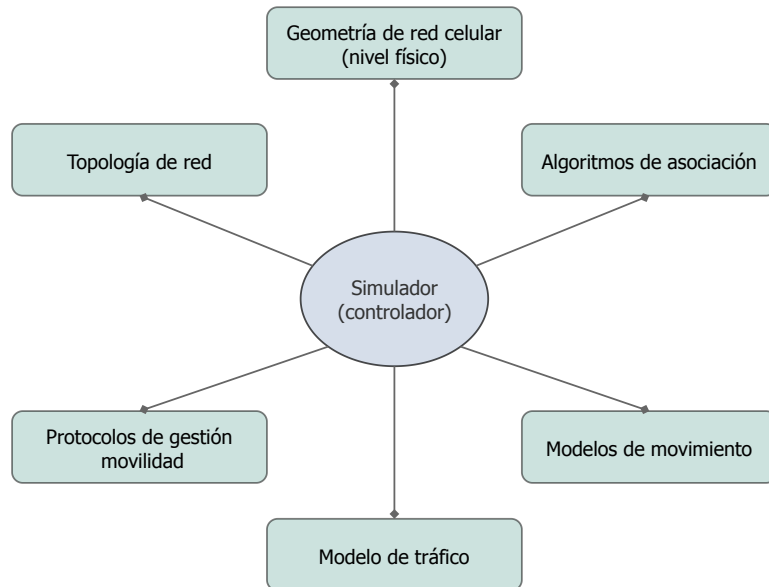


Figura 7.4: Diseño modular del simulador PyMMSim

El núcleo central de PyMMSim se encarga de controlar y gestionar la simulación de los eventos discretos que han sido considerados: el proceso de *handover*, el inicio de una demanda de tráfico y la finalización de la misma. Estos eventos dependen de la topología de la red de acceso seleccionada, del modelo de tráfico utilizado y del modelo de movimiento llevado a cabo por los nodos móviles en el escenario de red, entre otros. Además, tal y como se analizó en el Capítulo 5, la geometría de la red celular y los algoritmos de asociación entre la red de acceso y las estaciones base (nivel físico) también juegan un papel importante e influyen significativamente en los protocolos de gestión de la movilidad.

Por lo tanto, la operación del núcleo central del simulador se centra en la generación, planificación y procesamiento de los eventos discretos proporcionados por los módulos que componen el simulador. Tal y como se comentó anteriormente, los eventos discretos considerados pueden ser clasificados en dos tipos: sesiones o demandas de tráfico y movimientos de los nodos móviles. Por ello, el estado global de la simulación es actualizado solamente cuando ocurre alguno de estos eventos. Cada evento tiene un *timestamp* asociado que especifica cuándo ocurre el evento en cuestión, con el objetivo de permitir la ordenación de los mismos para su posterior tratamiento. Esta cola de eventos es generada y ordenada antes de comenzar la simulación, utilizando los módulos que componen PyMMSim.

Una vez se tiene inicializada y ordenada la cola de eventos, se lleva a cabo el proceso de simulación tal y como se muestra en la Figura 7.5.

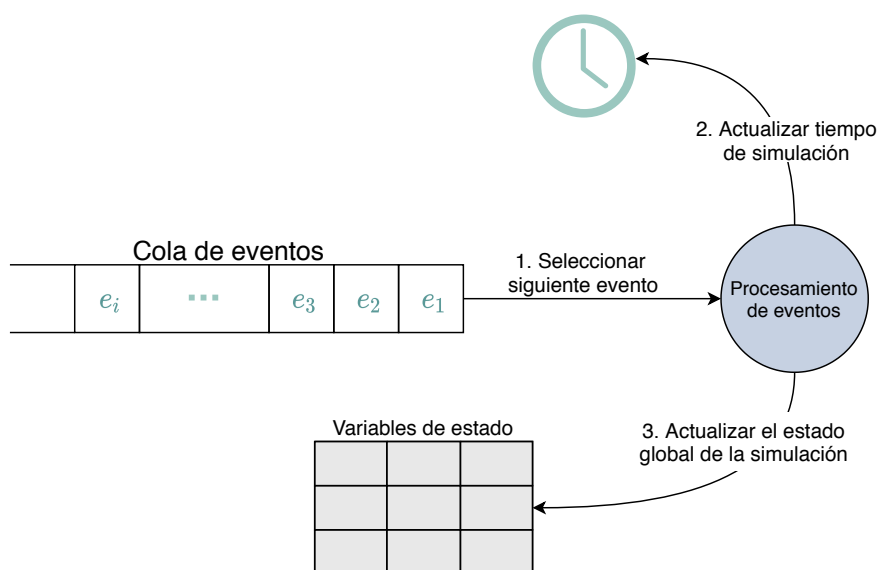


Figura 7.5: Simulación de eventos discretos con PyMMSim

El procesamiento se lleva a cabo de la siguiente manera: en primer lugar se selecciona el primer evento de la cola; posteriormente, se actualiza el tiempo de simulación hasta el *timestamp* del evento en cuestión y, finalmente, se procesa el evento como tal, actualizando las variables de estado global de la simulación.

Una vez explicado el funcionamiento del núcleo central de PyMMSim, se describen los módulos (Figura 7.4) de los que hace uso para llevar a cabo la simulación completa.

Modelo de red

Este módulo se encarga de modelar la topología de red utilizada en las simulaciones. PyMMSim considera una red de acceso representada como un grafo no dirigido $G = (V, E)$, donde V y E denotan los conjuntos de nodos y enlaces respectivamente. Para su implementación se ha utilizado como base una librería de Python llamada NetworkX [124], que permite la creación y manipulación de estructuras de redes complejas como grafos. Esta librería ha sido adaptada, introduciendo nuevas características y atributos a los nodos y enlaces que componen el grafo, con el objetivo de modelar por completo la red de acceso y tener en todo momento una visión del estado global de la red.

En definitiva, se ha utilizado el objeto *Graph* de NetworkX, considerando dos tipos de nodos:

- Nodos de interior: este tipo de nodos son nodos básicos de la librería NetworkX, a los que se les ha añadido como atributo una lista de flujos activos para tener constancia en todo momento de la carga del nodo en cuestión.
- Nodos de borde: estos nodos, además de contar con la lista de flujos activos, cuentan con una estructura de datos que almacena las estaciones base que él mismo controla. Tiene esta nueva funcionalidad porque es un nodo de borde y constituye el enlace entre la red de acceso y el nivel físico.

En cuanto a los enlaces entre los nodos, también se han realizado adaptaciones, con el objetivo de almacenar los flujos (demandas) que están pasando por un enlace en un momento determinado, pudiendo ofrecer en todo momento una carga real del sistema.

Por último, se han implementado algoritmos que actúan sobre nuestra topología de red, entre los que destacan el cálculo de caminos mínimos entre nodos de la propia topología y el cálculo de la centralidad entre nodos del grafo que se utiliza en el algoritmo de asignación LNA propuesto en el Capítulo 5.

Geometría de red celular

La red celular es modelada haciendo uso de geometría estocástica [148], donde una de las teorías más importantes es la relacionada con los procesos de puntos [149]. Inicialmente, su desarrollo fue estimulado por aplicaciones en biología, astronomía y ciencias de los materiales. En la actualidad, los procesos de puntos están siendo ampliamente utilizados en el contexto de las redes de comunicación [150].

Visualmente, un proceso de puntos puede ser descrito como una colección de puntos aleatorios en un área determinada. Desde un punto de vista formal, un proceso de puntos (PP) consiste en la representación de un espacio de probabilidad en un determinado espacio de medidas puntuales y finitas (puntos), es decir, se lleva a cabo una extrapolación de la distribución de probabilidad a un espacio de dos dimensiones, generando puntos con dos coordenadas (X e Y) [149, 150].

En definitiva, la geometría de red celular es modelada sobre una región de una determinada superficie, donde las estaciones base están distribuidas siguiendo procesos de puntos. El simulador PyMMSim cuenta con la implementación de dos procesos de puntos bien diferenciados, tal y como se muestra en la Figura 7.6.

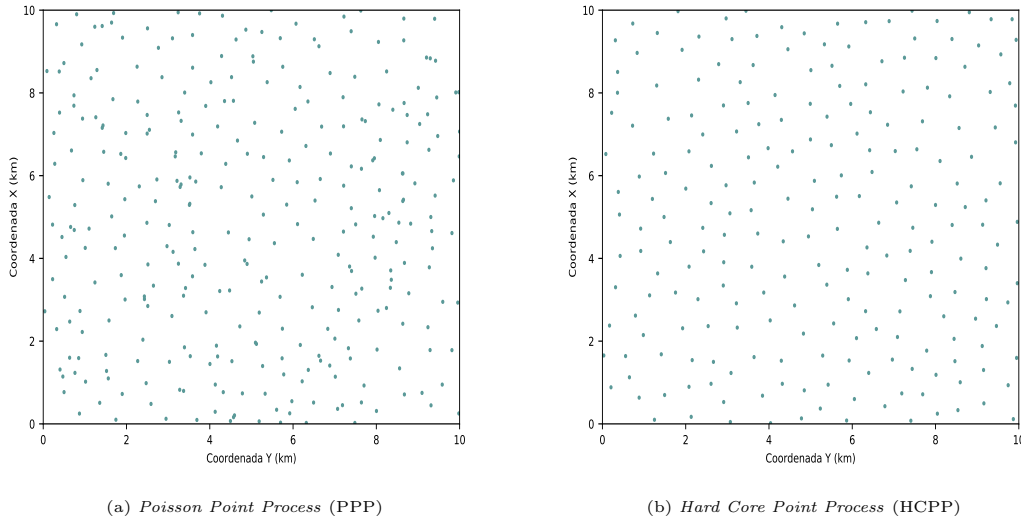


Figura 7.6: Procesos de puntos implementados en PyMMSim

En primer lugar (Figura 7.6a), se ha implementado el proceso *Poisson Point Process* (PPP) [151], cuya intensidad (λ_{BS}) coincide con el número medio de estaciones base (N_{BS}) por unidad de área (A) [125], y es obtenido como $\lambda_{BS} = N_{BS}/A$. Sin embargo, este modelo cuenta con una limitación y es que, debido a la independencia del propio

proceso, en algunos casos los puntos se ubican muy juntos. Por ello, se ha incorporado al simulador un proceso de puntos repulsivo con restricciones (Figura 7.6b), llamado *Hard Core Point Process* (HCPP) [150, 152], donde los puntos del proceso tienen prohibido estar más cerca que una cierta distancia mínima r_{min} [148, 153], conocida como parámetro *hard core*.

Por tanto, estos puntos representan la distribución de las estaciones base en PyMMSim. Además, el área de cobertura de estas estaciones base es modelada como una teselación de Poisson-Voronoi [126], en un espacio bidimensional donde cada usuario móvil se conecta a la estación base más cercana. La Figura 7.7 muestra cómo se ha modelado el área de cobertura de las estaciones base en PyMMSim.

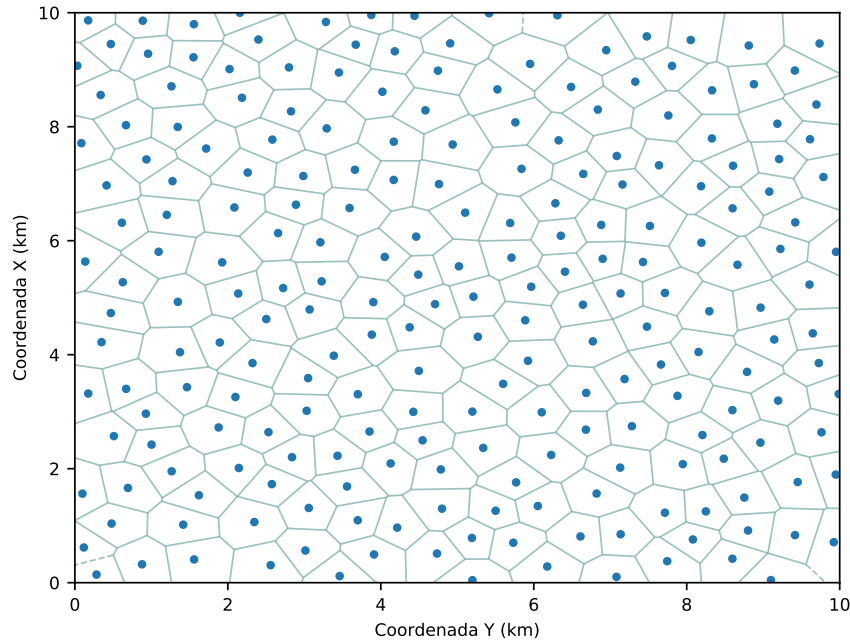


Figura 7.7: Área de cobertura de las estaciones base en PyMMSim según una teselación de Voronoi

Algoritmos de asociación entre la red de acceso y las estaciones base

Este módulo se encarga de llevar a cabo la asociación entre la red de acceso y las estaciones base generadas con los módulos anteriores. Para ello, hemos desarrollado una serie de algoritmos que fueron explicados en detalle en el Capítulo 5 y que se resumen a continuación:

- Algoritmo propuesto LNA: recopila la información de la topología de la red de acceso y examina la distribución de las estaciones base para realizar una asignación eficiente, consiguiendo importantes mejoras de rendimiento en el plano de control y en el plano de datos de los protocolos de gestión de la movilidad IP.
- Algoritmo K-Means: divide el conjunto de estaciones base en $|K|$ *clústeres* y asocia dichos *clústeres* a los *routers* de acceso de manera aleatoria.
- Enfoque *Balanced Closeness*: selecciona K estaciones base y calcula las N estaciones base más cercanas a ella, siendo $N = |B|/|K|$. Posteriormente, el algoritmo construye $|K|$ grupos que son asociados a los *routers* de acceso.

Modelos de movimiento

Este módulo se encarga de modelar la movilidad de los usuarios en el escenario de red propuesto por PyMMSim, permitiéndonos analizar el rendimiento de arquitecturas de red frente a los distintos comportamientos de los usuarios móviles. Los modelos de movilidad se diseñan con el objetivo de describir los patrones de movimiento de los nodos móviles, la posición en el escenario así como la velocidad y aceleración con la que cambian después de un cierto instante de tiempo [154]. Un modelo de movilidad debe ser lo más realista posible y, por tanto, implica el desarrollo de modelos complejos, ya que el movimiento humano, generalmente, no puede ser descrito de manera uniforme, sino que es un movimiento aleatorio que tiene que ser caracterizado convenientemente [155].

Para la implementación de estos modelos de movimiento se ha tomado como base una librería de Python llamada Pymobility [156], llevando a cabo una integración completa sobre PyMMSim, que actualmente cuenta con los siguientes modelos de movilidad:

- Random Walk [157].
- Random Waypoint [127].
- Truncated Levy Walk [158].
- Gauss-Markov [159].
- Reference Point Group Mobility model [160].
- Time-variant Community [161].

Protocolos de gestión de la movilidad

Este módulo se encarga del diseño e implementación de los protocolos de gestión de la movilidad IP utilizados por los nodos móviles durante la simulación. Actualmente, PyMMSim cuenta con la implementación de protocolos de gestión de la movilidad centralizada (MIPv6 y PMIPv6) y protocolos de gestión de la movilidad distribuida (NB-DMM), que son los que han sido evaluados durante el desarrollo de la presente Tesis, y que fueron explicados en detalle en el Capítulo 2.

Modelo de tráfico

Este módulo modela las sesiones de tráfico (demandas) de los nodos móviles, así como la duración de las mismas. Para su implementación, se ha tenido en cuenta lo explicado en capítulos anteriores: el número de sesiones entrantes a un MN sigue un proceso de Poisson con una tasa media λ_S , es decir, el tiempo entre llegadas de sesiones está exponencialmente distribuido con dicha tasa [46, 78, 106]. Además, es asumido que la duración de una sesión está exponencialmente distribuida con parámetro μ_S [107, 85]. La Figura 7.8 muestra la distribución de demandas (d_t) en el tiempo de simulación, contemplando los eventos asociados a ellas que son considerados por el simulador PyMMSim: el inicio (i_t) y el final (f_t) de las demandas de tráfico.

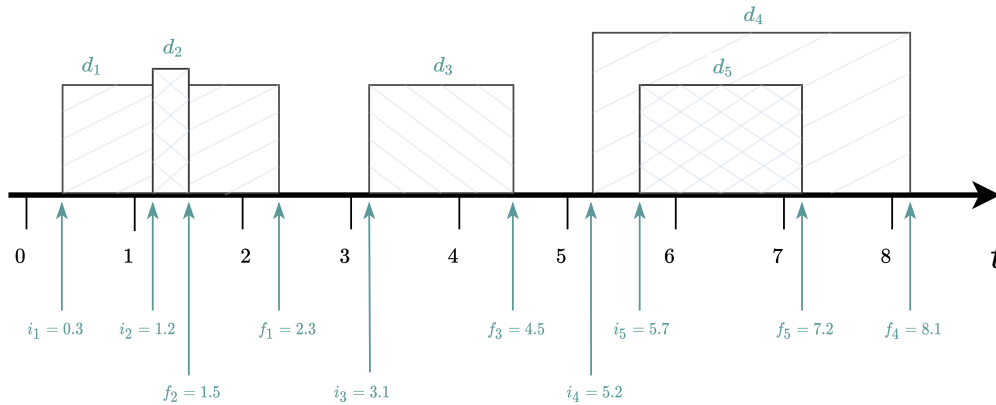


Figura 7.8: Distribución de las demandas en el tiempo de simulación

Finalmente, relacionado con el modelo de tráfico y para aportar más realismo a PyMMSim, se ha añadido una funcionalidad en la que el flujo o tasa de datos de las demandas varíen entre 1500 Kbps y 10 Mbps [129].

7.3. Escenario de simulación

En esta sección se describe un escenario de simulación genérico usado para evaluar el rendimiento de los protocolos y propuestas de gestión de la movilidad, así como los distintos algoritmos de asignación entre la red de acceso y las estaciones base. Hay que tener en cuenta que todos los componentes del escenario de red definidos por PyMMSim son parametrizables. A continuación se detalla uno de los escenarios de red más utilizados en esta Tesis.

En primer lugar, en cuanto a las topologías de la red de acceso, PyMMSim cuenta con la implementación de distintos tipos de topologías, entre las que destacan las que se muestran en la Figura 7.9. Estas topologías se basan en las topologías de red utilizadas en [129], variando el número de *routers* de acceso.

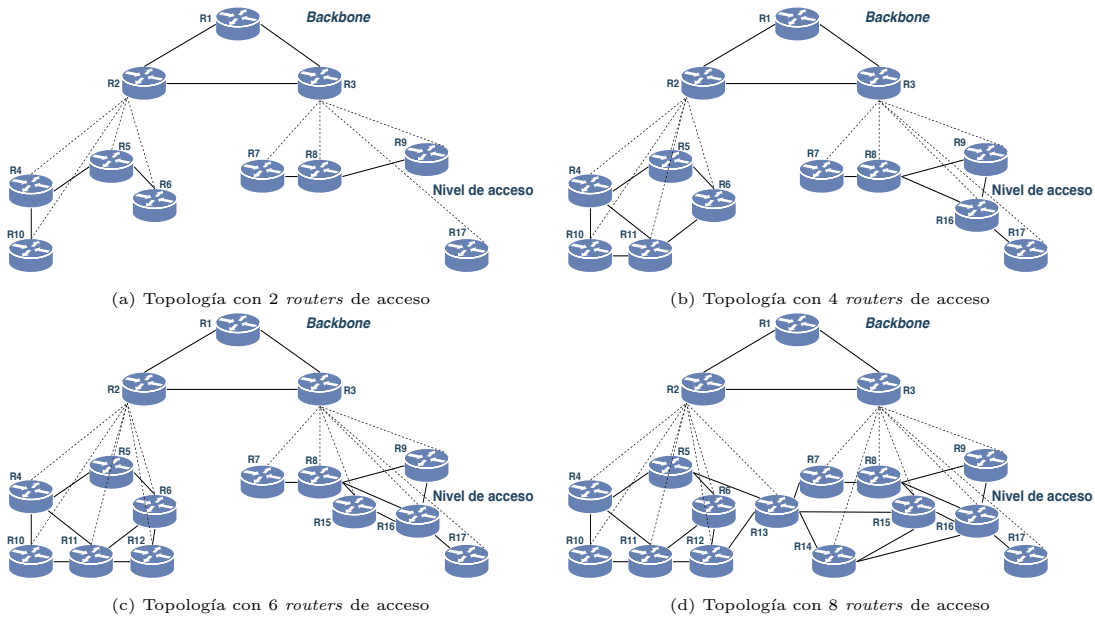


Figura 7.9: Topologías de red implementadas en PyMMSim

Además, el escenario de simulación está formado por una región cuadrada de $10 \times 10 \text{ km}^2$, donde las estaciones base están distribuidas siguiendo un proceso de puntos llamado *Hard Core Point Process* (HCPP) con parámetro $r = 0.5$. Además, el área de cobertura de las estaciones base es modelada como una teselación de Poisson-Voronoi [126], en un espacio bidimensional donde cada usuario móvil se conecta a la estación base más cercana.

El movimiento de los usuarios es definido a través del modelo de movilidad Random Waypoint [127], con una velocidad uniformemente distribuida entre 1 y 20 m/s . Cada

simulación normalmente está formada por 200 usuarios móviles que se mueven por el dominio de movilidad, conectándose a diferentes estaciones base. Estos usuarios móviles gestionan un conjunto de sesiones durante el tiempo de simulación, asumiéndose que el número de sesiones entrantes por usuario móvil sigue un proceso de Poisson con una tasa media $\lambda = 0.01$, y la duración de una sesión está exponencialmente distribuida con parámetro $\mu = 10$ [128]. También, se debe tener en cuenta que el flujo o tasa de datos de una demanda varía entre 1500 Kbps y 10 Mbps (por ejemplo, *stream* de vídeo) [129].

El Algoritmo 5 muestra el procedimiento seguido por PyMMSim desde la asociación entre la red de acceso y la geometría de red celular hasta la obtención de los resultados finales.

Algoritmo 5 Operación del núcleo de simulación en PyMMSim.

Entrada: Topología de red NT , geometría de red celular RC y lista de nodos móviles $listaMN$

Salida: Resultados globales de la simulación

▷ **Asociación de la red de acceso y las estaciones base**

1: $NT = \text{assocNetPhy}(NT, RC);$

▷ **Unificación y ordenación de todos los eventos de simulación**

2: $listaEventos = [];$

3: **for** $mn_i \in listaMN$ **do**

4: $listaEventos = listaEventos + mn_i.\text{getLisEvents}();$

5: **end for**

6: $listaEventos = \text{sorted}(listaEventos);$

▷ **Simulación de cada evento**

7: **for** $e_i \in listaEventos$ **do**

8: $mn = listaMN[e_i.\text{getIdMN}];$

9: **if** e_i es *handover* **then**

10: $NT = mn.\text{handoverProcess}(NT, e_i);$

11: **else if** e_i es *demanda* **then**

12: $NT = mn.\text{treatDemand}(NT, e_i);$

13: **end if**

14: **end for**

▷ **Obtención de los resultados finales de la simulación**

15: $resultados = NT.\text{getSimulationResults}();$

16: **return** $resultados$

7.4. Entorno de ejecución

En la última década se ha producido una gran evolución de las arquitecturas de computación de altas prestaciones (HPC), que ha sido bien recibida en el contexto de la simulación, permitiendo a los investigadores llevar a cabo análisis computacionales con mayor fiabilidad en el menor tiempo posible [162].

Por ello, PyMMSim se ha integrado en un entorno real de HPC, permitiendo realizar simulaciones con topologías de red a gran escala, obteniendo resultados más fiables gracias al despliegue masivo del simulador sobre un *clúster* de supercomputación distribuido, formado por un gran número de máquinas. La infraestructura utilizada ha sido la del supercomputador LUSITANIA II, descrita en el Apéndice A. Para el correcto funcionamiento de PyMMSim en el entorno HPC ha sido necesario diseñar una estructura específica, tal y como se muestra en la Figura 7.10.

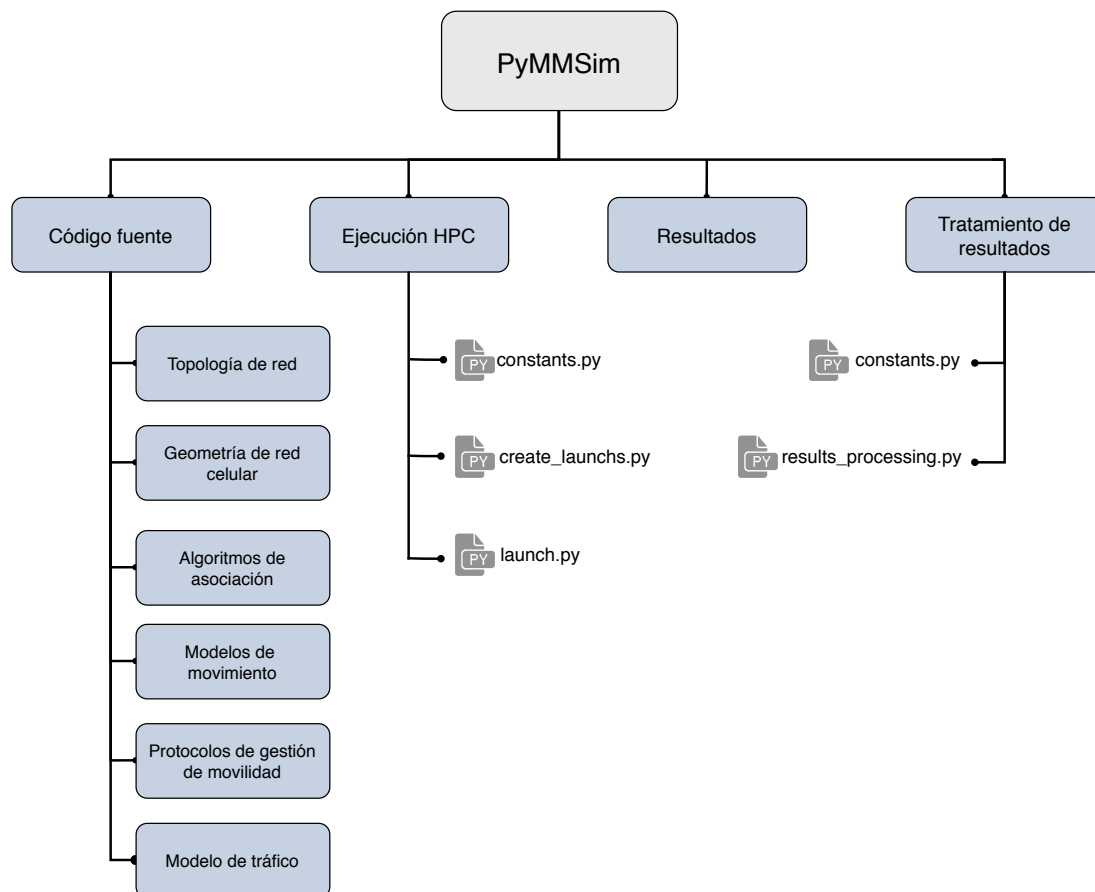


Figura 7.10: Integración de PyMMSim en un entorno HPC

La estructura de PyMMSim en el supercomputador está formada por los siguientes elementos:

- Código fuente. Contiene la implementación completa del simulador desarrollado PyMMSim, con todos los módulos que lo componen.
- Ejecución HPC. El supercomputador LUSITANIA II cuenta con un gestor de recursos y tareas (SLURM, [163]) para controlar todos los procesos que se ejecuten sobre él. SLURM funciona en dos fases: primero se crea un fichero que contiene todos los parámetros necesarios para la ejecución y, posteriormente, se lanza este fichero al gestor de colas. Por ello, hemos incorporado a PyMMSim la funcionalidad de automatizar el proceso de creación de ficheros necesarios para lanzar las ejecuciones masivas; además de automatizar el propio proceso de lanzamiento de estos ficheros sobre SLURM.
- Resultados. Es un directorio con una estructura específica que almacena todos los resultados de la simulación en formato JSON, que son analizados posteriormente. La estructura de este directorio será explicada más adelante.
- Tratamiento de resultados. Debido a que PyMMSim lleva a cabo ejecuciones de simulación de manera masiva, es necesario disponer de un módulo de tratamiento de resultados que automatice el proceso. Por ello, hemos incorporado esta funcionalidad a PyMMSim. Este módulo trata los resultados de la simulación, calculando medidas estadísticas que son necesarias para representarlos, utilizando librerías como Scipy y Matplotlib.

En definitiva, se ha implementado un entorno de pruebas que nos ha permitido ejecutar una simulación de Montecarlo, donde en cada iteración el tiempo de simulación es de 1000 segundos, proporcionando los valores medios y mejorando la precisión de los resultados con un intervalo de confianza del 95 %. A continuación se detalla el procedimiento que se ha seguido para ejecutar en el entorno HPC:

- Configuración de los ficheros de constantes (*constants.py*) del módulo *Ejecución HPC*. En este fichero se configura el número de nodos móviles que actúan en cada conjunto de pruebas, la cola de SLURM a la cual se lanzarán las ejecuciones, el directorio base donde se encuentra el código fuente de PyMMSim y el número de experimentos que se lanzarán por cada conjunto de nodos móviles.
- Ejecutar el fichero *create_launchs.py* para crear las carpetas donde se almacenarán los ficheros que se lanzarán sobre SLURM (fichero *run*) y los resultados asociados a cada prueba (fichero JSON).

- Ejecutar el fichero *launch.py*, que realiza el lanzamiento de las pruebas al supercomputador.

Una vez llevado a cabo este procedimiento, se tendrá una estructura de directorios en el módulo *Resultados*, tal y como la que se muestra en la Figura 7.11. Como se puede observar, se realizan pruebas con distintos conjuntos de nodos móviles (desde 1 hasta 100). A su vez, para cada conjunto de MN las pruebas se repiten 50 veces (*test1MN_1* ... *test1MN_50*), con el objetivo de obtener mayor fiabilidad en los resultados.

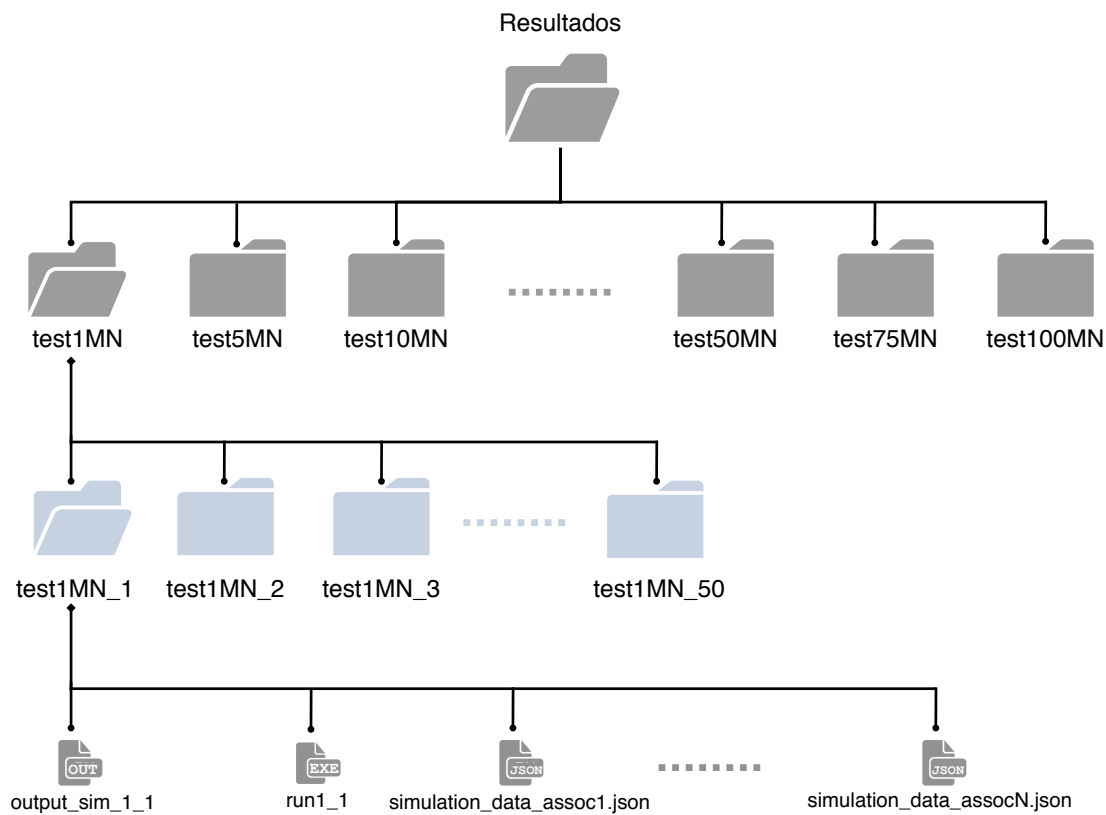


Figura 7.11: Estructura del directorio de resultados de PyMMSim

Finalmente, dentro de cada una de estas carpetas se tiene la siguiente información:

- Fichero con la salida estándar de PyMMSim: *output_sim*.
- Fichero de ejecución de cada prueba, que será lanzado al supercomputador a través del gestor de colas SLURM.
- Ficheros con los resultados globales de la simulación para cada asociación entre la red de acceso y las estaciones base. Los distintos algoritmos de asociación

utilizados por PyMMSim fueron explicados en la Sección 7.2. El simulador utiliza el formato JSON para reportar los resultados globales de la simulación.

Una vez que se tienen los resultados globales hay que ejecutar el fichero *results_processing.py* del módulo *Tratamiento de resultados*. Este fichero funciona en tres fases:

1. Procesamiento: se lleva a cabo una unificación de los resultados para cada conjunto de nodos móviles. Para su explicación recurriremos a la Figura 7.11: el fichero accede a los resultados de las simulaciones para cada conjunto de nodos móviles, es decir, si estamos tratando las simulaciones de 1 nodo móvil, se accede a la carpeta *test1MN* y a su vez, en cada subcarpeta (*test1MN_1 ... test1MN_50*), se trata cada fichero JSON de cada asociación por separado. De esta manera, se unifican todos los resultados de un conjunto de nodos móviles para una asociación determinada en un solo fichero llamado *sim_results_assoc1.json*.
2. Cálculos estadísticos: se accede a cada fichero generado en la fase anterior y se calcula la media y el intervalo de confianza, almacenando los resultados en un fichero llamado *total_results_mean_conf_int_assoc1.json*.
3. Generación de gráficas: a partir de los resultados de la fase anterior, se generan las gráficas de resultados con Matplotlib.

7.5. Métricas para evaluar la gestión de la movilidad

En esta sección se explican las métricas de rendimiento introducidas en PyMMSim para evaluar la eficiencia de los protocolos de gestión de la movilidad desde dos puntos de vista:

- Plano de control. El rendimiento del plano de control es evaluado a través de la sobrecarga introducida por los protocolos de movilidad en cuestión, lo que se conoce como coste de señalización. PyMMSim calcula este valor para cada nodo móvil y, al terminar la simulación, calcula el coste de señalización global introducido por todos los nodos móviles durante todo el tiempo de simulación.
- Plano de datos. En el caso del plano de datos, el rendimiento es evaluado desde dos perspectivas diferentes:
 - Coste de entrega de paquetes. Es la métrica más relevante para medir el rendimiento del plano de datos en los protocolos de gestión de la movilidad

IP. Este valor está influenciado por el tamaño medio de los mensajes de datos multiplicado por el número de saltos necesarios para reenviar los paquetes desde el CN hasta el MN, y viceversa.

- Coste de tunelización. Es una métrica similar al coste de entrega de paquetes, pero únicamente representa el coste de añadir el *overhead* del mecanismo de encapsulación para enviar la información por el túnel.

Estas métricas son ampliamente utilizadas en los análisis y evaluaciones de soluciones de movilidad [78, 85, 88, 93, 101], permitiendo determinar el rendimiento global de la red móvil.

Además, el simulador PyMMSim tiene la capacidad de calcular los costes económicos asociados a la distribución de antenas propuesta en cada simulación. Los costes económicos considerados [131, 132] fueron explicados en el Capítulo 6 y se resumen a continuación:

- Costes de capital (CAPEX). También llamados inversiones en bienes de capitales. El CAPEX está íntimamente ligado a los costes de la infraestructura fija de la empresa y se amortizan con el tiempo. Para un operador de red, el CAPEX incluye los costes relacionados con la compra de terrenos y/o edificios (por ejemplo, para alojar el equipamiento tecnológico), la infraestructura de red y el despliegue inicial de dicha infraestructura.
- Costes de operación (OPEX). También llamados costes de funcionamiento o costes operativos. El OPEX no contribuye a los costes del despliegue de la infraestructura; representan el coste necesario para mantener operativa la infraestructura, incluyendo los costes de operaciones técnicas y comerciales, administración, etc. Para un operador de red, el OPEX está formado por los costes de alquiler (edificio, terreno, equipos de red, etc), los costes de energía y los costes de mantenimiento, administración y operación, entre otros.

Finalmente, a modo de resumen, la Tabla 7.1 muestra las métricas utilizadas por PyMMSim y el ámbito de actuación de cada una de ellas.

Tabla 7.1: Métricas para la evaluación de la gestión de la movilidad en PyMMSim

	Plano de Control	Plano de Datos	Costes Económicos
Costes de señalización	✓		
Costes de entrega de paquetes		✓	
Costes de tunelización		✓	
Costes de capital			✓
Costes de operación			✓

7.6. Conclusiones

En esta capítulo se ha descrito el diseño y la implementación de un simulador de gestión de la movilidad en redes de nueva generación (PyMMSim) desarrollado en Python, que ha sido utilizado para obtener parte de los resultados presentados en esta Tesis.

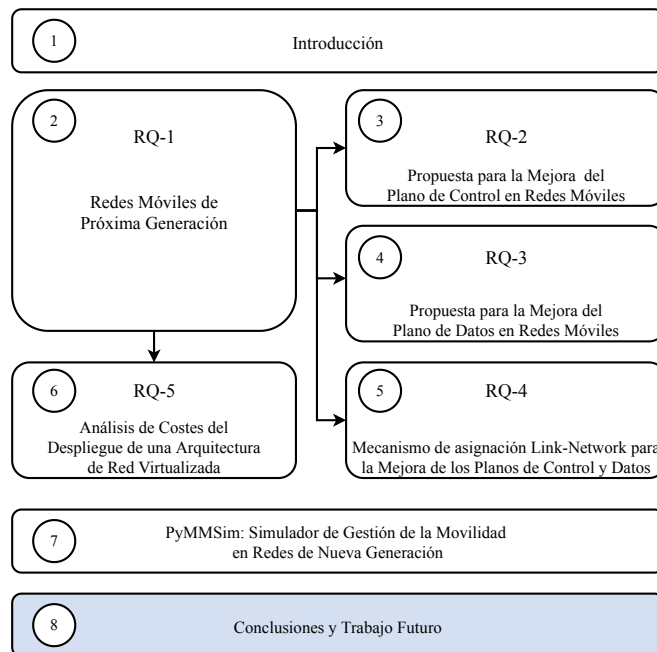
PyMMSim simplifica la generación de resultados de movilidad, permitiendo evaluar la escalabilidad y rendimiento de diferentes protocolos de gestión de la movilidad y las propuestas desarrolladas a lo largo de la Tesis. Su diseño modular ofrece la posibilidad de realizar ampliaciones, añadiendo nuevas funcionalidades fácilmente.

Además, PyMMSim se ha integrado en un entorno de computación de altas prestaciones, permitiéndonos realizar simulaciones de topologías de red a gran escala y obteniendo resultados más fiables.

Capítulo 8

Conclusiones y Trabajo Futuro

En este capítulo se describen las conclusiones y aportaciones principales, junto con las líneas de trabajo futuras que surgen a partir de las investigaciones realizadas en esta Tesis.



8.1. Conclusiones

En los últimos años, el incremento exponencial del tráfico de datos móviles, unido a la proliferación de aplicaciones que hacen uso de los propios servicios de datos móviles, han propiciado que los operadores de red actualicen sus infraestructuras y

busquen nuevos mecanismos que permitan una gestión eficiente de la red, de acuerdo a las demandas específicas de los usuarios. Como resultado, los operadores, junto a la comunidad científica, están investigando y diseñando mecanismos y estrategias para enfrentarse a la nueva realidad propuesta por la tecnología 5G, con el objetivo de ser capaces de gestionar los recursos de red de forma dinámica y eficiente.

Todos estos cambios implican una adaptación de las redes de acceso y de los protocolos de comunicaciones utilizados hasta el momento. Dentro de estas actuaciones hay que tener en cuenta uno de los procesos más importantes involucrados en la gestión de la red, el soporte a la movilidad, que engloba el mantenimiento de la conectividad entre diferentes redes de acceso y la gestión de los recursos requeridos por los usuarios móviles. En los últimos años, las organizaciones más relevantes en el campo de las comunicaciones móviles (IETF y 3GPP) han diseñado diferentes soluciones de gestión de la movilidad. Sin embargo, estas soluciones no presentan comportamientos adecuados en ciertos escenarios de movilidad de nueva generación. Por ello, en esta Tesis se proponen una serie de mecanismos novedosos con el objetivo de mejorar el rendimiento de los planos de control y de datos, desde la perspectiva de los protocolos de gestión de la movilidad.

En primer lugar, se ha proporcionado una visión general de las redes móviles de próxima generación, describiendo la evolución que han experimentado las arquitecturas de red en los últimos años, y prestando especial atención al proceso de soporte a la movilidad. Por ello, se han analizado los principales enfoques en este contexto, así como su evolución, desde los enfoques de red centralizados a los enfoques distribuidos, introduciendo la necesidad de nuevos mecanismos que solucionen los problemas y limitaciones que afectan en la actualidad a dichos protocolos de gestión de la movilidad. Este análisis previo motivó en gran medida parte de la investigación llevada a cabo en esta Tesis, en la cual se proponen tres nuevos mecanismos para mejorar el rendimiento de las redes móviles desde tres perspectivas diferentes.

Nuestra primera propuesta, TE-DMM, permite llevar a cabo una gestión eficiente del plano de control en arquitecturas distribuidas, consiguiendo una disminución importante del tráfico de control. Se ha proporcionado una evaluación de rendimiento analítica y experimental entre la propuesta TE-DMM y la solución más representativa de gestión de la movilidad distribuida (NB-DMM). El modelo analítico ha permitido comparar ambas propuestas, poniendo de manifiesto los beneficios introducidos por TE-DMM en términos de costes de señalización, los cuales son reducidos aproximadamente en un 50 % con respecto a NB-DMM. Además, el análisis revela que TE-DMM reduce la complejidad del plano de control, mejorando el proceso de gestión de túneles durante el movimiento de los usuarios. Por otro lado, para realizar

la evaluación experimental se han utilizado implementaciones reales de NB-DMM y nuestro mecanismo (TE-DMM) sobre un *testbed* de movilidad real. Los resultados obtenidos muestran que TE-DMM mejora la latencia de *handover*, proporcionando, además, beneficios significativos en relación a la pérdida de paquetes durante dicho proceso.

La segunda propuesta, SR-DMM, combina SDN con DMM para mejorar el proceso de gestión de la movilidad desde el punto de vista del plano de datos, proporcionando flexibilidad, escalabilidad y fiabilidad a las redes móviles de próxima generación. La propuesta ha sido evaluada analítica y experimentalmente. El modelo analítico revela los beneficios que aporta SR-DMM al rendimiento global de la red en términos de coste de señalización y coste de entrega de paquetes. Estas mejoras son conseguidas debido a que SR-DMM reduce la complejidad del plano de datos y la gestión de túneles IP sobre IP, evitando el uso de los mismos durante el movimiento de los usuarios móviles. Por otro lado, para realizar la evaluación experimental se ha llevado a cabo la implementación de nuestra propuesta sobre el controlador SDN Ryu, permitiendo así realizar una evaluación experimental utilizando implementaciones reales sobre un *testbed* de movilidad real. Los resultados demuestran que los mecanismos propios de SDN introducen una latencia mínima en el proceso de *handover*.

Nuestro tercer mecanismo propone una estrategia de asociación entre estaciones base y la red de acceso, llamada LNA, que proporciona beneficios al rendimiento de los protocolos de gestión de la movilidad, desde dos puntos de vista: plano de control y plano de datos. Este mecanismo es evaluado y comparado con respecto a otros en términos de costes de movilidad: costes de señalización y costes de entrega de paquetes. En primer lugar, se ha formulado un problema de optimización para minimizar el impacto de la asignación de estaciones base a los *routers* de acceso. Además, se ha desarrollado un nuevo algoritmo (LNA), con el objetivo de llevar a cabo una asignación adecuada mediante la recopilación de información de la topología de la red de acceso y el análisis de la distribución de las estaciones base en un área determinada. Finalmente, se lleva a cabo una evaluación de rendimiento del algoritmo propuesto con respecto a otros algoritmos de asignación, en términos de costes de señalización y costes de entrega de paquetes, permitiendo evaluar el rendimiento global de la red móvil. Los resultados obtenidos demuestran que nuestro algoritmo LNA puede reducir con éxito los costes de señalización hasta un 35 % en comparación con los algoritmos de referencia, sin penalizar el coste de entrega de paquetes que también se mejora, aunque en menor medida. Esta reducción de ambas métricas es lo que hace realmente interesante a la propuesta LNA, ya que las soluciones de gestión de la movilidad tienen como objetivo balancear la sobrecarga de señalización y los costes asociados al plano de datos, pero, tal

y como se ha discutido en capítulos anteriores, generalmente la disminución de uno de los costes asociados a uno de los planos impacta negativamente en el otro, y viceversa.

Además de los tres mecanismos propuestos y explicados anteriormente, se ha diseñado e implementado un simulador de gestión de la movilidad en redes de nueva generación (PyMMSim) desarrollado en Python, que ha sido utilizado para obtener los resultados de simulación presentados en esta Tesis. PyMMSim simplifica la generación de resultados de movilidad, permitiendo evaluar la escalabilidad y rendimiento de diferentes protocolos de gestión de la movilidad y las propuestas desarrolladas a lo largo de la Tesis. Su diseño modular ofrece la posibilidad de realizar ampliaciones, añadiendo fácilmente nuevas funcionalidades. Además, PyMMSim se ha integrado en un entorno de computación de altas prestaciones, permitiéndonos realizar simulaciones de topologías de red a gran escala y obteniendo resultados más fiables.

Finalmente, se han evaluado los costes relacionados con el despliegue de tecnologías de red emergentes como SDN y NFV, analizando la viabilidad económica y proponiendo un modelo de costes basado en costes de capital (CAPEX) y costes de operación (OPEX). De esta manera, hemos obtenido una evaluación comparativa fiable entre arquitecturas de red tradicionales y arquitecturas de red virtualizadas, en la que se revela que la arquitectura C-RAN proporciona importantes ahorros de CAPEX (70 %), OPEX (59 %) y TCO (68 %).

8.2. Trabajo futuro

Tras el desarrollo de los diferentes mecanismos propuestos durante la investigación llevada a cabo en esta Tesis, han surgido algunas líneas de trabajo adicionales que necesitan ser tratadas en profundidad en trabajos futuros, ya que se han considerado fuera del alcance de la Tesis.

- Optimización de la red móvil y de los mecanismos de gestión de la movilidad para soportar eficientemente servicios y aplicaciones 5G, sobre redes virtuales de próxima generación. En este contexto, sería interesante optimizar la ubicación de las funciones virtuales de red teniendo en cuenta la movilidad de los usuarios, prestando especial atención a la manera en que se ofrecen servicios 5G a través de SFC (*Service Function Chaining*). En entornos de *network slicing*, la movilidad de usuario tiene que ser gestionada, no solamente, entre diferentes estaciones base o tecnologías de acceso, sino también entre diferentes *slices*; motivando la necesidad de nuevas soluciones de gestión de la movilidad que permitan a los usuarios móviles mantener las sesiones activas entre *slices*.

- Preparación y optimización del proceso de handover entre diferentes *slices* a través del desarrollo de nuevos mecanismos para predecir el deterioro o pérdida de rendimiento de las condiciones de las diferentes *slices*, con el objetivo de convertir el proceso de *handover* en un procedimiento proactivo. En este sentido, se puede aprovechar el potencial de la analítica de datos y las técnicas de *machine learning*.
- Caracterizar y modelar el canal de propagación de redes 5G ultradensas, integrándolo en nuestro simulador PyMMSim, para ofrecer simulaciones integrales (*cross-layer*) extremo a extremo. De esta manera se podrá analizar el impacto de los diferentes parámetros que componen el nivel físico sobre el plano de control y el plano de datos de las soluciones de gestión de la movilidad IP; tomando como referencia los principales casos de uso propuestos en entornos 5G. En relación al simulador PyMMSim también se podría contemplar la integración de otros simuladores centrados en el nivel físico (red de acceso radio) de 5G.
- Analizar y combinar nuevos modelos estocásticos de geometría de red celular con los modelos propuestos de nivel de red, realizando las simulaciones necesarias sobre un entorno de computación de altas prestaciones, con el objetivo de detectar parámetros que tengan cierto impacto en los planos de control y de datos de los protocolos de gestión de la movilidad IP.
- Desarrollar nuevas estrategias de asociación entre la red de acceso y el nivel físico, a través de la implementación de nuevos algoritmos basados en técnicas de *clustering* que consideren métricas adicionales a la distancia; analizando el rendimiento del plano de control y del plano de datos en los protocolos de gestión de la movilidad IP y optimizando la complejidad computacional de las propuestas desarrolladas. Además, sería interesante estudiar el impacto de estas estrategias de asociación sobre parámetros que juegan un papel importante en redes móviles de nueva generación, como son la calidad de servicio (QoS) y la calidad de experiencia (QoE).
- Aplicar mecanismos de optimización matemática en el modelo de costes basado en CAPEX y OPEX propuesto en esta Tesis, con el objetivo de ajustar parámetros críticos considerados por los operadores de red. Estudios recientes [164] han estimado que se producirá una disminución del gasto medio por usuario itinerante en Europa. Básicamente, esto implica que cualquier coste relacionado con el CAPEX y OPEX derivado del aumento de tráfico y motivado por la entrada en vigor del RLAH (*Roam Like At Home*) no generará ingresos a los operadores de red. Además, el aumento de tráfico de datos móviles conlleva asociado inversiones de costes de capital en el PGW [165]. Por ello, es necesario estudiar y analizar la

manera de reducir los costes operacionales y de infraestructura de los operadores de red eficientemente, intentando conseguir un equilibrio entre los propios costes y la infraestructura desplegada.

- Integración de tecnologías de gestión de red habilitadoras de 5G (NFV y MEC) sobre el simulador PyMMSim, analizando la relación entre las propias tecnologías y el impacto de ellas sobre el proceso de gestión de la movilidad; prestando especial atención y analizando los principales casos de uso de 5G que han sido establecidos por el 5G PPP. Además, en este contexto, sería interesante analizar nuevos *datasets* de trazas reales de tráfico y movilidad de usuarios, integrándolo en nuestro simulador PyMMSim. De esta manera, se podrían comparar los diferentes mecanismos propuestos en esta Tesis en diferentes áreas de actuación (rurales, urbanas, vehiculares, etc).

Apéndice A

Infraestructura Utilizada en las Simulaciones

En este apéndice se presenta la infraestructura usada para obtener los resultados presentados en esta Tesis. Dado que se han llevado a cabo evaluaciones y análisis de resultados que requieren de una alta capacidad de cómputo, ha sido necesario el uso de un sistema de computación de altas prestaciones. Por ello, se ha utilizado una parte del supercomputador LUSITANIA II, gestionado por CénitS (Centro Extremeño de iNvestigación, Innovación Tecnológica y Supercomputación).

Las características técnicas de la infraestructura utilizada en CénitS se detallan a continuación.

Nodos de cómputo utilizados en las simulaciones y evaluaciones

- 40 servidores Fujitsu Primergy CX2550 con 2 procesadores Intel Xeon E5-2660v3, de 10 cores cada uno, a 2,6 GHz (20 cores por nodo, 800 cores en total) y 25 MB de caché, con 80 GB de RAM y 2 discos SSD 128 GB.
- 168 IBM System x iDataPlex dx360 M4 con 2 procesadores Intel E5-2670 SandyBridge-EP, de 8 cores cada uno, a 2.6 GHz (16 cores por nodo, 2688 cores en total) 20 MB de caché y 32 GB de RAM.

Sistema de almacenamiento de altas prestaciones

- Sistema de ficheros paralelo y distribuido para computación de altas prestaciones (Lustre) formado por:

- Cabina de datos (OST) Eternus DX 200 (41 discos de 2TB NL-SAS y 31 discos de 4TB NL-SAS) = 206 TB.
- Cabina de metadatos (MDT) Eternus DX 200S3 (15 discos de 900 GB SAS) = 12 TB.
- 2 servidores Fujitsu Primergy RX2530 M1 con 2 procesadores Intel Xeon E5-2620v3 (6 cores a 2,4 GHz y 15 MB de caché); 64 GB de RAM DDR4 y 2 discos SAS de 300 GB para gestión de metadatos con Lustre.
- 2 servidores Fujitsu Primergy RX2530 M1 con 2 procesadores Intel Xeon E5-2620v3 (6 cores a 2,4 GHz y 15 MB de caché); 64 GB de RAM DDR4, 2 discos SAS de 300 GB para gestión de los objetos con Lustre.

Red de comunicaciones

- La conectividad del supercomputador con el exterior se resuelve a través de una conexión de 10 Gbps con la Red Científico Tecnológica de Extremadura, que conecta las principales ciudades y centros tecnológicos de la región. Interconectada a su vez con RedIRIS y con la red europea GÉANT.
- La infraestructura de cálculo se vertebra sobre una red formada por:
 - 14 *switches* de baja latencia Infiniband Mellanox SX6036 de 36 puertos FDR a 56 Gbps.
 - 3 *switches* de baja latencia InfiniBand Mellanox IS5030 de 36 puertos QDR a 40 Gbps.

Referencias

- [1] R. Li, Z. Zhao, X. Zhou, G. Ding, Y. Chen, Z. Wang, and H. Zhang, “Intelligent 5G: When Cellular Networks Meet Artificial Intelligence,” *IEEE Wireless Communications*, vol. 24, pp. 175–183, October 2017.
- [2] 3GPP, “System Architecture for the 5G System, V15.4.0, 3GPP Standard 23.501,” December 2018.
- [3] I. F. Akyildiz, S. Nie, S.-C. Lin, and M. Chandrasekaran, “5G Roadmap: 10 Key enabling Technologies,” *Computer Networks*, vol. 106, pp. 17 – 48, 2016.
- [4] B. Blanco, J. O. Fajardo, I. Giannoulakis, E. Kafetzakis, S. Peng, J. Pérez-Romero, I. Trajkovska, P. S. Khodashenas, L. Goratti, M. Paolino, E. Sfakianakis, F. Liberal, and G. Xilouris, “Technology Pillars in the Architecture of Future 5G Mobile Networks: NFV, MEC and SDN,” *Computer Standards & Interfaces*, vol. 54, pp. 216–228, 2017.
- [5] R. Buyya and S. N. Srirama, *Management and Orchestration of Network Slices in 5G, Fog, Edge, and Clouds*, pp. 79–101. Wiley, 2019.
- [6] Internet Engineering Task Force (IETF), <https://www.ietf.org/>.
- [7] Third Generation Partnership Project (3GPP), <https://www.3gpp.org/>.
- [8] S. Jeon, S. Figueiredo, R. L. Aguiar, and H. Choo, “Distributed Mobility Management for the Future Mobile Networks: A Comprehensive Analysis of Key Design Options,” *IEEE Access*, vol. 5, pp. 11423–11436, 2017.
- [9] J. Carmona-Murillo, I. Soto, F.-J. Rodríguez-Pérez, D. Cortés-Polo and J.-L. González-Sánchez, “Performance Evaluation of Distributed Mobility Management Protocols: Limitations and Solutions for Future Mobile Networks,” *Mobile Information Systems*, vol. 2017, 2017.

- [10] Y. Mehmood, N. Haider, M. Imran, A. Timm-Giel, and M. Guizani, "M2M Communications in 5G: State-of-the-Art Architecture, Recent Advances, and Research Challenges," *IEEE Communications Magazine*, vol. 55, pp. 194–201, September 2017.
- [11] S. Li, L. D. Xu, and S. Zhao, "5G Internet of Things: A survey," *Journal of Industrial Information Integration*, vol. 10, pp. 1 – 9, 2018.
- [12] B. Kitchenham, "Procedures for Performing Systematic Reviews," *Keele, UK, Keele Univ.*, vol. 33, August 2004.
- [13] K. BA and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," vol. 2, January 2007.
- [14] A. Fink, "Conducting Research Literature Reviews: From the Internet to Paper," 1998.
- [15] IEEE Digital Library, <https://ieeexplore.ieee.org/>.
- [16] Scopus, <https://www.scopus.com/>.
- [17] Parsifal, <https://parsif.al/>.
- [18] Cisco Systems Inc., "Cisco Visual Networking Index: Forecast and Trends, 2017-2022, White Paper," February 2019.
- [19] Ericsson, "Ericsson Mobility Report: The Future of Mobile Subscriptions," November 2019.
- [20] 3GPP, "Service requirements for next generation new services and markets, 3GPP TS 22.261, Version 15.5.0, Release 15.," July 2018.
- [21] E. L. A. Jain and I. Demirkol, "Mobility Management as a Service for 5G Networks," in *Proceedings of the 14th International Symposium on Wireless Communication Systems*, pp. 28–31, 2017.
- [22] H. Ji, S. Park, J. Yeo, Y. Kim, J. Lee, and B. Shim, "Ultra-Reliable and Low-Latency Communications in 5G Downlink: Physical Layer Aspects," *IEEE Wireless Communications*, vol. 25, pp. 124–130, June 2018.
- [23] J. Martín-Pérez, L. Cominardi, C. J. Bernardos, A. De la Oliva, and A. Azcorra, "Modeling Mobile Edge Computing Deployments for Low Latency Multimedia Services," *IEEE Transactions on Broadcasting*, vol. 65, pp. 464–474, June 2019.

- [24] N. Akkari and N. Dimitriou, “Mobility Management Solutions for 5G Networks: Architecture and Services,” *Computer Networks*, vol. 169, p. 107082, 2020.
- [25] F. Z. Yousaf, M. Bredel, S. Schaller, and F. Schneider, “NFV and SDN—Key Technology Enablers for 5G Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 35, p. 2468–2478, November 2017.
- [26] S. Sezer, S. Scott-Hayward, P. K. Chouhan, B. Fraser, D. Lake, J. Finnegan, N. Viljoen, M. Miller, and N. Rao, “Are we ready for SDN? Implementation Challenges for Software-Defined Networks,” *IEEE Communications Magazine*, vol. 51, pp. 36–43, July 2013.
- [27] M. Jammal, T. Singh, A. Shami, R. Asal, and Y. Li, “Software-Defined Networking: State of the Art and Research Challenges,” *CoRR*, vol. abs/1406.0124, 2014.
- [28] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, “Software-Defined Networking: A Comprehensive Survey,” *Proceedings of the IEEE*, vol. 103, pp. 14–76, January 2015.
- [29] H. Kim and N. Feamster, “Improving network management with software defined networking,” *IEEE Communications Magazine*, vol. 51, pp. 114–119, February 2013.
- [30] B. A. A. Nunes, M. Mendonca, X. Nguyen, K. Obraczka, and T. Turletti, “A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks,” *IEEE Communications Surveys Tutorials*, vol. 16, pp. 1617–1634, March 2014.
- [31] Foundation Open Networking, “Software-Defined Networking: The new norm for networks,” *ONF White Paper*, vol. 2, pp. 2–6, 2012.
- [32] R. Trivisonno, R. Guerzoni, I. Vaishnavi, and D. Soldani, “SDN-based 5G Mobile Networks: architecture, functions, procedures and backward compatibility,” *Transactions on Emerging Telecommunications Technologies*, vol. 26, no. 1, pp. 82–92, 2015.
- [33] Foundation Open Networking, “Open flow switch specification,” 2016.
- [34] A. Doria, J. H. Salim, R. Haas, H. Khosravi, W. Wang, L. Dong, R. Gopal, and J. Halpern, “Forwarding and Control Element Separation (ForCES) Protocol Specification,” *Internet Requests for Comments, RFC Editor, RFC*, vol. 5810, 2010.

- [35] ETSI Industry Specification Group (ISG) NFV, “ETSI GS NFV-MAN 001 V1.1.1: Network Functions Virtualization (NFV): Management and Orchestration,” 2014.
- [36] Y. Li and M. Chen, “Software-Defined Network Function Virtualization: A Survey,” *IEEE Access*, vol. 3, pp. 2542–2553, 2015.
- [37] R. Mijumbi, J. Serrat, J. Gorricho, N. Bouten, F. De Turck, and R. Boutaba, “Network Function Virtualization: State-of-the-Art and Research Challenges,” *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 236–262, 2016.
- [38] B. Han, V. Gopalakrishnan, L. Ji, and S. Lee, “Network Function Virtualization: Challenges and Opportunities for Innovations,” *IEEE Communications Magazine*, vol. 53, pp. 90–97, February 2015.
- [39] J. A. Cabrera, R. Schmoll, G. T. Nguyen, S. Pandi, and F. H. P. Fitzek, “Softwarization and Network Coding in the Mobile Edge Cloud for the Tactile Internet,” *Proceedings of the IEEE*, vol. 107, pp. 350–363, February 2019.
- [40] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, “Mobile Edge Computing: a key technology towards 5G,” *White Paper*, vol. 11, 2015.
- [41] P. Mach and Z. Becvar, “Mobile edge computing: A survey on architecture and computation offloading,” *IEEE Communications Surveys Tutorials*, vol. 19, no. 3, pp. 1628–1656, 2017.
- [42] A. Zafeiropoulos, A. Gavras, A. Tzanakaki, A. Albanese, A. Kousaridas, A. Weit, B. Sayadi, B. Jou, C. Bernardos, C. Benzaid, C. Mannweiler, D. Camps-Mur, D. Breitgand, D. Estevez, D. Navratil, D. Mi, D. Lopez, D. Klonidis, E. Mutafulungwa, E. Fotopoulou, E. Kafetzakis, E. Pateromichelakis, E. Biton, F. Tesema, G. Kalfas, H. Karl, J. Bartelt, J. Gutiérrez, J. Cosmas, J. Thomson, J. Giménez, J. Alcaraz Calero, J. Mangues-Bafalluy, K. Katsalis, L. Gallo, M. Gramaglia, M. Spada, M. Salih, N. Nikaein, N. Jawad, N. Maletic, Ö. Bulakci, P. Demestichas, P. Hasselmeyer, Q. Wang, Q. Wei, R. Ustok, R. Blom, S. Pontarelli, S. Keskin, S. Salsano, S. Parker, T. Deiss, U. Acar, X. Li, and Y. Zhang, *5G PPP Architecture Working Group: View on 5G Architecture*, vol. Version 3.0. Belgium: European Commission, June 2019.
- [43] M. M. Sajjad, D. Jayalath, and C. J. Bernardos, “A Comprehensive Review of Enhancements and Prospects of Fast Handovers for Mobile IPv6 Protocol,” *IEEE Access*, vol. 7, pp. 4948–4978, 2019.

- [44] A. Jain, E. Lopez-Aguilera, and I. Demirkol, “Evolutionary 4G/5G Network Architecture Assisted Efficient Handover Signaling,” *IEEE Access*, vol. 7, pp. 256–283, 2019.
- [45] K. Kong, W. Lee, Y. Han, M. Shin, and H. You, “Mobility Management for all-IP Mobile Networks: Mobile IPv6 vs. Proxy Mobile IPv6,” *IEEE Wireless Communications*, vol. 15, no. 2, pp. 36–45, 2008.
- [46] J. Carmona-Murillo, V. Friderikos, and J. González-Sánchez, “A hybrid dmm solution and trade-off analysis for future wireless networks,” *Computer Networks*, vol. 133, pp. 17 – 32, 2018.
- [47] T.-X. Do and Y. Kim, “Control and data plane separation architecture for supporting multicast listeners over distributed mobility management,” *ICT Express*, vol. 3, no. 2, pp. 90 – 95, 2017. Special Issue on Patents, Standardization and Open Problems in ICT Practices.
- [48] N. Feamster, J. Rexford, and E. Zegura, “The road to sdn: An intellectual history of programmable networks,” *SIGCOMM Comput. Commun. Rev.*, vol. 44, p. 87–98, April 2014.
- [49] L. Yang, R. Dantu, T. Aderson, and R. Gopal, “Forwarding and Control Element Separation (ForCES) Framework,” *Internet Engineering Task Force*, April 2004.
- [50] J. Salim, H. Khosravi, A. Kleen, and A. Kuznetsov, “Linux Netling as an IP Services Protocol,” *Internet Engineering Task Force, RFC 3549*, July 2003.
- [51] M. Caesar, D. Caldwell, N. Feamster, J. Rexford, A. Shaikh, and J. van der Merwe, “Design and implementation of a routing control platform,” in *Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation*, NSDI’05, (Berkeley, CA, USA), pp. 15–28, USENIX Association, 2005.
- [52] N. Feamster, H. Balakrishnan, J. Rexford, A. Shaikh, and J. van der Merwe, “The case for separating routing from routers,” in *Proceedings of the ACM SIGCOMM Workshop on Future Directions in Network Architecture*, (New York, NY, USA), p. 5–12, Association for Computing Machinery, 2004.
- [53] T. V. Lakshman, T. Nandagopal, R. Ramjee, K. Sabnani, and T. Woo, “The SoftRouter Architecture,” in *In ACM HOTNETS*, 2004.
- [54] A. Farrel, J. Vasseur, and J. Ash, “A Path Computation Element (PCE)-Based Architecture,” *Internet Engineering Task Force, RFC 4655*, August 2006.

- [55] J. M. Smith, K. L. Calvert, S. L. Murphy, H. K. Orman, and L. L. Peterson, "Activating networks: a progress report," *Computer*, vol. 32, no. 4, pp. 32–41, 1999.
- [56] J. Biswas, A. A. Lazar, J. . Huard, Koonseng Lim, S. Mahjoub, L. . Pau, M. Suzuki, S. Torstensson, Weiguo Wang, and S. Weinstein, "The IEEE P1520 standards initiative for programmable network interfaces," *IEEE Communications Magazine*, vol. 36, no. 10, pp. 64–70, 1998.
- [57] J. E. van der Merwe, S. Rooney, L. Leslie, and S. Crosby, "The tempest-a practical framework for network programmability," *IEEE Network*, vol. 12, no. 3, pp. 20–28, 1998.
- [58] J. Van der Merwe, A. Cepleanu, K. D'Souza, B. Freeman, A. Greenberg, D. Knight, R. McMillan, D. Moloney, J. Mulligan, H. Nguyen, M. Nguyen, A. Ramarajan, S. Saad, M. Satterlee, T. Spencer, D. Toll, and S. Zelingher, "Dynamic Connectivity Management with an Intelligent Route Service Control Point," in *Proceedings of the 2006 SIGCOMM Workshop on Internet Network Management*, (New York, NY, USA), p. 29–34, Association for Computing Machinery, 2006.
- [59] M. Casado, M. Freedman, J. Pettit, J. Luo, N. McKeown, and S. Shenker, "ETHANE: Taking Control of the Enterprise," *Computer Communication Review - CCR*, vol. 37, pp. 1–12, 10 2007.
- [60] H. Yan, D. A. Maltz, H. Gogineni, and Z. Cai, "Tesseract: A 4D Network Control Plane ," in *4th USENIX Symposium on Networked Systems Design & Implementation (NSDI 07)*, (Cambridge, MA), USENIX Association, April 2007.
- [61] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "OpenFlow: Enabling Innovation in Campus Networks," *SIGCOMM Comput. Commun. Rev.*, vol. 38, p. 69–74, March 2008.
- [62] N. Gude, T. Koponen, J. Pettit, B. Pfaff, M. Casado, N. McKeown, and S. Shenker, "NOX: Towards an Operating System for Networks," *SIGCOMM Comput. Commun. Rev.*, vol. 38, p. 105–110, July 2008.
- [63] Open vSwitch, <https://www.openvswitch.org/>.
- [64] B. Pfaff, J. Pettit, T. Koponen, K. Amidon, M. Casado, and S. Shenker, "Extending networking into the virtualization layer," in *8th ACM Workshop on Hot Topics in Networks (HotNets-VIII)*, October 2009.

- [65] ONOS: Open network operating system, <https://www.opennetworking.org/onos/>.
- [66] 3GPP, “Architecture enhancements for control and user plane separation of EPC nodes, 3GPP TS 23.214, Release 14,” June 2017.
- [67] C. Perkins, D. Johnson, and J. Akko, “Mobility Support in IPv6. RFC 6275,” July 2011.
- [68] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, “Proxy Mobile IPv6. RFC 5213,” August 2008.
- [69] H. Chan, D. Liu, P. Seite, H. Yokota, and J. Korhonen, “Requirements for distributed mobility management. RFC7333,” August 2014.
- [70] T. Condeixa and S. Sargento, “Centralized, Distributed or Replicated IP Mobility?,” *IEEE Communications Letters*, vol. 18, pp. 376–379, February 2014.
- [71] E. Kaljic, A. Maric, P. Njemcevic, and M. Hadzialic, “A Survey on Data Plane Flexibility and Programmability in Software-Defined Networking,” *IEEE Access*, vol. 7, pp. 47804–47840, 2019.
- [72] K. Benzekki, A. El Fergougui, and A. Elbelrhiti Elalaoui, “Software-defined networking (SDN): a survey,” *Security and Communication Networks*, vol. 9, no. 18, pp. 5803–5833, February 2016.
- [73] D. Liu, J. C. Zuniga, P. Seite, H. Chan, and C. J. Bernardos, “Distributed Mobility Management: Current practices and gap analysis. RFC 7429.,” January 2015.
- [74] P.P. Ernest, H.A. Chan, O. E. Falowo, and L.A. Magagula, “Distributed Mobility Management with Distributed Routing Management at Access Routers for Network-Based Mobility Support,” *Wireless Personal Communications*, vol. 84, pp. 181–205, 2015.
- [75] J. Lee, J. Bonnin, and X. Lagrange, “Host-based distributed mobility management support protocol for ipv6 mobile networks,” in *2012 IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 61–68, 2012.
- [76] J. Lee, J. Bonnin, I. You, and T. Chung, “Comparative Handover Performance Analysis of IPv6 Mobility Management Protocols,” *IEEE Transactions on Industrial Electronics*, vol. 60, no. 3, pp. 1077–1088, 2013.

- [77] M.-S. Kim, S. Lee, D. Cypher, and N. Golmie, "Performance analysis of fast handover for proxy mobile ipv6," *Information Sciences*, vol. 219, pp. 208 – 224, 2013.
- [78] H. Ali-Ahmad, M. Ouzzif, P. Bertin, and X. Lagrange, "Performance Analysis on Network-Based Distributed Mobility Management," *Wireless Personal Communications*, vol. 74, no. 4, p. 1245–1263, 2014.
- [79] D. Shin, D. Moses, M. Venkatachalam, and S. Bagchi, "Distributed mobility management for efficient video delivery over all-IP mobile networks: Competing approaches," *IEEE Network*, vol. 27, no. 2, pp. 28–33, 2013.
- [80] J. Calle-Cancho, J. Mendoza-Rubio, J. L. González-Sánchez, D. Cortés-Polo, and J. Carmona-Murillo, "Towards a software-based mobility management for 5g: An experimental approach for flattened network architectures," *Computer Science and Information Systems*, vol. 17, no. 1, pp. 51–70, 2020.
- [81] C. Yu, J. Lan, Z. Guo, Y. Hu, and T. Baker, "An Adaptive and Lightweight Update Mechanism for SDN," *IEEE Access*, vol. 7, pp. 12914–12927, 2019.
- [82] Y. Yang, S. Jeon, D. Kim, and H. Choo, "Design and implementation of sdn-based lightweight ipv6 mobility manager in onos," in *Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication*, IMCOM '17, (New York, NY, USA), Association for Computing Machinery, 2017.
- [83] P. Berde, M. Gerola, J. Hart, Y. Higuchi, M. Kobayashi, T. Koide, B. Lantz, B. O'Connor, P. Radoslavov, W. Snow, and G. Parulkar, "ONOS: towards an open, distributed SDN OS," *HotSDN 2014 - Proceedings of the ACM SIGCOMM 2014 Workshop on Hot Topics in Software Defined Networking*, 08 2014.
- [84] Haneul Ko, Insun Jang, Jaewook Lee, Sangheon Pack, and Giwon Lee, "SDN-based distributed mobility management for 5G," in *2017 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 116–117, 2017.
- [85] C. Makaya and S. Pierre, "An Analytical Framework for Performance Evaluation of IPv6-Based mobility Management Protocols," *Wireless Communications, IEEE Transactions on*, vol. 7, pp. 972 – 983, 04 2008.
- [86] J. Lee, T. Ernst, and T. Chung, "Cost analysis of IP mobility management protocols for consumer mobile devices," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 1010–1017, 2010.

- [87] J. Lee, J. Bonnin, P. Seite, and H. A. Chan, "Distributed IP mobility management from the perspective of the IETF: motivations, requirements, approaches, comparison, and challenges," *IEEE Wireless Communications*, vol. 20, no. 5, pp. 159–168, 2013.
- [88] F. Giust, C. J. Bernardos, and A. de la Oliva, "Analytic Evaluation and Experimental Validation of a Network-Based IPv6 Distributed Mobility Management Solution," *IEEE Transactions on Mobile Computing*, vol. 13, no. 11, pp. 2484–2497, 2014.
- [89] K. Vasu, S. Mahapatra, and K. Cheruvu, "A Comprehensive Framework for Evaluating IPv6 Based Mobility Management Protocols," *Wireless Personal Communications*, vol. 78, pp. 943–977, September 2014.
- [90] K. Munir, X. Lagrange, P. Bertin, K. Guillouard, and M. Ouzzif, "Performance analysis of mobility management architectures in cellular networks," *Telecommunication Systems*, vol. 59, June 2015.
- [91] S. Jeon, N. Kang, D. Corujo, and R. Aguiar, "Comprehensive Performance Evaluation of Distributed and Dynamic Mobility Routing Strategy," *Computer Networks*, vol. 79, January 2015.
- [92] P.-W. Park, S.-M. Kim, and S.-G. Min, "OpenFlow-Based Mobility Management Scheme and Data Structure for the Mobility Service at Software Defined Networking," *International Journal of Distributed Sensor Networks*, vol. 2016, pp. 1–9, March 2016.
- [93] P. P. Ernest, O. E. Falowo, and H. A. Chan, "Design and performance evaluation of distributed mobility management schemes for network mobility," *Journal of Network and Computer Applications*, vol. 61, pp. 46 – 58, 2016.
- [94] S. M. Raza, D. S. Kim, D. Shin, and H. Choo, "Leveraging proxy mobile IPv6 with SDN," *Journal of Communications and Networks*, vol. 18, no. 3, pp. 460–475, 2016.
- [95] M. I. Sanchez, A. de la Oliva, and V. Mancuso, "Experimental Evaluation of an SDN-Based Distributed Mobility Management Solution," in *Proceedings of the Workshop on Mobility in the Evolving Internet Architecture*, MobiArch '16, (New York, NY, USA), p. 31–36, Association for Computing Machinery, 2016.
- [96] T. Nguyen, C. Bonnet, and J. Harri, "SDN-based distributed mobility management for 5G networks," in *2016 IEEE Wireless Communications and Networking Conference*, pp. 1–7, 2016.

- [97] M. Murtadha, N. Noordin, B. Ali, and F. Hashim, "Design and simulation analysis of network-based fully distributed mobility management in flattened network architecture," *Telecommunication Systems*, vol. 65, September 2017.
- [98] L. Cominardi, F. Giust, C. J. Bernardos, and A. D. L. Oliva], "Distributed mobility management solutions for next mobile network architectures," *Computer Networks*, vol. 121, pp. 124 – 136, 2017.
- [99] C.-M. Huang, T. Dao Duy, and M.-S. Chiang, "A Bursty Multi-node Handover scheme for mobile internet using the partially Distributed Mobility Management (BMH-DMM) architecture," *Telecommunication Systems*, vol. 69, February 2018.
- [100] M. Aman, S. Mahfooz, M. Zubair, N. Mukhtar, K. Imran, and S. Khusro, "Tunnel-Free Distributed Mobility Management (DMM) Support Protocol for Future Mobile Networks," *Electronics* 2019, 8, 1519.
- [101] E. M. O. Fafolahan and S. Pierre, "A Seamless Mobility Management Protocol in 5G Locator Identifier Split Dense Small Cells," *IEEE Transactions on Mobile Computing*, pp. 1–1, 2019.
- [102] J. Guan, V. Sharma, I. You, M. Atiquzzaman, and M. Imran, "Extension of MIH for FPMIPv6 (EMIH-FPMIPv6) to support optimized heterogeneous handover," *Future Generation Computer Systems*, vol. 97, March 2019.
- [103] J. Calle-Cancho, D. Cortés-Polo, J. L. González-Sánchez, L. I. Jiménez, and J. Carmona-Murillo, "Tunnel-Free Distributed Mobility Management (DMM) Support Protocol for Future Mobile Networks," *Electronics* 2020, 9, 417.
- [104] M. Balfaqih, Z. Balfaqih, V. Shepelev, S. Alharbi, and W. Al-Areeqi, "An analytical framework for distributed and centralized mobility management protocols," *Journal of Ambient Intelligence and Humanized Computing*, February 2020.
- [105] T. Nguyen and C. Bonnet, "A hybrid centralized-Distributed Mobility Management for supporting highly mobile users," in *2015 IEEE International Conference on Communications (ICC)*, pp. 3945–3951, 2015.
- [106] L. Kleinrock, "Queueing Systems: Theory. Wiley: New York, NY, USA, 1975.," 1975.
- [107] M. Murtadha, N. Noordin, B. Ali, and F. Hashim, "Design and evaluation of distributed and dynamic mobility management approach based on PMIPv6 and MIH protocols," *Wireless Networks*, vol. 21, April 2015.

- [108] I. Akyildiz and W. Wang, “A dynamic location management scheme for next-generation multitier PCS systems,” *Wireless Communications, IEEE Transactions on*, vol. 1, pp. 178 – 189, February 2002.
- [109] G. Wan and E. Lin, “Cost Reduction in Location Management Using Semi-realtime Movement Information,” *Wireless Networks*, vol. 5, July 1998.
- [110] D. Cortés-Polo, J. González-Sánchez, J. Carmona-Murillo, and F. Rodríguez-Pérez, “Proposal and analysis of integrated ptn architecture in the mobile backhaul to improve the qos of hetnets,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, December 2015.
- [111] MAD-PMIPv6 Project’s Website, “Open Platform for Distributed Mobility Management Solutions (ODMM),” <http://odmm.net>.
- [112] C. J. Bernardos, A. De la Oliva, and F. Giust, “A PMIPv6-Based Solution for Distributed Mobility Management,” *draft-bernardos-dmm-pmip-09. IMDEA Networks Institute: Madrid, Spain*.
- [113] CORE, “Common Open Research Emulator,” <https://www.nrl.navy.mil/itd/ncs/products/core>.
- [114] I. U. Din, S. Hassan, M. K. Khan, M. Guizani, O. Ghazali, and A. Habbal, “Caching in Information-Centric Networking: Strategies, Challenges, and Future Research Directions,” *IEEE Communications Surveys Tutorials*, vol. 20, no. 2, pp. 1443–1474, 2018.
- [115] O. A. Khan, M. A. Shah, I. Ud Din, B. Kim, H. A. Khattak, J. J. P. C. Rodrigues, H. Farman, and B. Jan, “Leveraging Named Data Networking for Fragmented Networks in Smart Metropolitan Cities,” *IEEE Access*, vol. 6, pp. 75899–75911, 2018.
- [116] M. He, A. Martínez Alba, A. Basta, A. Blenk, and W. Kellerer, “Flexibility in Softwarized Networks: Classifications and Research Challenges,” *IEEE Communications Surveys & Tutorials*, pp. 1–1, January 2019.
- [117] W. A. Simpson, D. T. Narten, E. Nordmark, and H. Soliman, “Neighbor Discovery for IP version 6 (IPv6).,” *Internet Engineering Task Force, RFC 4861*, September 2007.
- [118] D. Cortés-Polo, J. Calle-Cancho, J. Carmona-Murillo, and J. González-Sánchez, “Future Trends in Mobile-Fixed Integration for Next Generation Networks::

Classification and Analysis,” *International Journal of Vehicular Telematics and Infotainment Systems*, vol. 1, pp. 33–53, January 2017.

- [119] Ryu SDN Framework, <https://ryu-sdn.org/>.
- [120] M. Karimzadeh, H. Berg, R. de O. Schmidt, and A. Pras, “Quantitative Comparison of the Efficiency and Scalability of the Current and Future LTE Network Architectures,” *Wireless Communications and Mobile Computing*, vol. 2017, pp. 1–20, November 2017.
- [121] D. Arthur and S. Vassilvitskii, “k-means++: The advantages of careful seeding,” in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 61–68, 2007.
- [122] Linton C. Freeman, “Centrality in social networks conceptual clarification,” *Social Networks*, vol. 1, pp. 215–239, 1978.
- [123] G. Zheng, A. Tsiopoulos, and V. Friderikos, “Optimal VNF Chains Management for Proactive Caching,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6735–6748, 2018.
- [124] A. Hagberg, D. Schult and P. Swart, “Exploring network structure, dynamics, and function using networkx,” in *Proceedings of the 7th Python in Science conference (SciPy 2008)*, 2008.
- [125] M. Di Renzo, A. Zappone, T. T. Lam and M. Debbah, “System-Level Modeling and Optimization of the Energy Efficiency in Cellular Networks—A Stochastic Geometry Framework,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2539–2556, 2018.
- [126] M. Berg, M. Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*. January 2008.
- [127] E. Hyytia, P. Lassila, and J. Virtamo, “Spatial node distribution of the random waypoint mobility model with applications,” *IEEE Transactions on Mobile Computing*, vol. 5, no. 6, pp. 680–694, 2006.
- [128] M.K. Murtadha, N.K. Noordin, B.M. Ali and F. Hashim, “Design and evaluation of distributed and dynamic mobility management approach based on PMIPv6 and MIH protocols,” *Wireless Networks*, pp. 1–17, 2015.
- [129] G. Zheng, C. Wang, V. Friderikos and M. Dohler, “High mobility multi modal e-health services,” in *IEEE International Conference on Communications (ICC)*, 2018.

- [130] P.K. Agyapong, M. Iwamura, D. Staehle, W. Kiess, A. Benjebbour, "Design considerations for a 5G network architecture," *IEEE Communications Magazine*, pp. 65–75, 2014.
- [131] S. Verbrugge, D. Colle, M. Pickavet, P. Demeester, S. Pasqualini, A. Iselt, A. Kirstädter, R. Hülsermann, F.-J. Westphal, and M. Jaeger, "Methodology and input availability parameters for calculating OPEX and CAPEX costs for realistic network scenarios," *Journal of Optical Networking*, vol. 5, no. 6, pp. 509–520, 2006.
- [132] B. Naudts, M. Kind, S. Verbrugge, D. Colle, and M. Pickavet, "How can a mobile service provider reduce costs with software-defined networking?," *International Journal of Network Management*, vol. 26, no. 1, pp. 56–72, 2016.
- [133] W. H. Chin, Z. Fan, and R. Haines, "Emerging technologies and research challenges for 5G wireless networks," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 106–112, 2014.
- [134] I. Akyildiz, p. Wang, and S.-C. Lin, "SoftAir: A software defined networking architecture for 5G wireless systems," *Computer Networks*, vol. 85, May 2015.
- [135] E. Hernandez-Valencia, S. Izzo, and B. Polonsky, "How will NFV/SDN transform service provider opex?," *IEEE Network*, vol. 29, no. 3, pp. 60–67, 2015.
- [136] N. Zhang and H. Hämmäinen, "Cost efficiency of sdn in lte-based mobile networks: Case finland," in *2015 International Conference and Workshops on Networked Systems (NetSys)*, pp. 1–5, 2015.
- [137] D. Pompili, A. Hajisami, and H. Viswanathan, "Dynamic Provisioning and Allocation in Cloud Radio Access Networks (C-RANs)," *Ad Hoc Networks*, vol. 30, March 2015.
- [138] M. M. Rahman, C. Despins, and S. Affes, "Analysis of CAPEX and OPEX benefits of wireless access virtualization," in *2013 IEEE International Conference on Communications Workshops (ICC)*, pp. 436–440, 2013.
- [139] S. F. Yunas, J. Niemelä, M. Valkama, and T. Isotalo, "Techno-economical analysis and comparison of legacy and ultra-dense small cell networks," in *39th Annual IEEE Conference on Local Computer Networks Workshops*, pp. 768–776, 2014.
- [140] K. Johansson, A. Furuskar, P. Karlsson, and J. Zander, "Relation between base station characteristics and cost structure in cellular systems," in *2004 IEEE 15th*

International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE Cat. No.04TH8754), vol. 4, pp. 2627–2631 Vol.4, 2004.

- [141] LTE Multi-Platform Base Station. [Online] June 2020. Available: <https://www.winncom.com/en/products/998-03-536>
- [142] CNMC. Información Geográfica de las estaciones base LTE de la Comunidad Autónoma de Extremadura.[Online] July 2020. Available: <http://data.cnmc.es/datagraph/jsp/graph/mapa.jsp>
- [143] A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis*. McGraw Hill, 1991.
- [144] K. Wehrle, M. Güneş, and J. Gross, *Modeling and Tools for Network Simulation*. Springer, 2010.
- [145] G. A. Wainer, *Discrete-Event Modeling and Simulation: A Practitioner's Approach*. Taylor & Francis Group, 2009.
- [146] J. Banks, J. S. Carson, B. L. Nelson, and D. M. Nicol, *Discrete-Event System Simulation*. Prentice Hall, 2005.
- [147] A. M. Law, *Simulation Modeling and Analysis*. McGraw Hill, 2007.
- [148] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*. John Wiley and Sons, 1996.
- [149] D. Daley and D. V. Jones, *An introduction to the theory of point processes*. Springer, 1988.
- [150] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, “Stochastic geometry and random graphs for the analysis and design of wireless networks,” *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, 2009.
- [151] J. G. Andrews, F. Baccelli, and R. K. Ganti, “A Tractable Approach to Coverage and Rate in Cellular Networks,” *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, 2011.
- [152] H. ElSawy and E. Hossain, “A Modified Hard Core Point Process for Analysis of Random CSMA Wireless Networks in General Fading Environments,” *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1520–1534, 2013.
- [153] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge University Press, 2012.

- [154] P. Santi, *Mobility Models for Next Generation Wireless Networks: Ad Hoc, Vehicular and Mesh Networks*. John Wiley and Sons, 2012.
- [155] M. M. Zonoozi and P. Dassanayake, “User mobility modeling and characterization of mobility patterns,” *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 7, pp. 1239–1252, 1997.
- [156] A. Panisson, *Pymobility*. <https://github.com/panisson/pymobility>.
- [157] K. Pearson, “The Problem of the Random Walk,” *Nature*, vol. 72, no. 1865, p. 294, 1905.
- [158] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, and S. Chong, “On the Levy-Walk Nature of Human Mobility,” *IEEE/ACM Transactions on Networking*, vol. 19, no. 3, pp. 630–643, 2011.
- [159] T. Camp, J. Boleng, and V. Davies, “A Survey of Mobility Models for Ad Hoc Network Research,” *Wireless Communications and Mobile Computing*, vol. 2, August 2002.
- [160] X. Hong, M. Gerla, G. Pei, and C.-C. Chiang, “A group mobility model for ad hoc wireless networks,” in *Proceedings of the 2nd ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, MSWiM ’99, (New York, NY, USA), p. 53–60, Association for Computing Machinery, 1999.
- [161] W. . Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, “Modeling Time-Variant User Mobility in Wireless Mobile Networks,” in *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*, pp. 758–766, 2007.
- [162] H. Casanova, A. Giersch, A. Legrand, M. Quinson, and F. Suter, “Versatile, scalable, and accurate simulation of distributed applications and platforms,” *Journal of Parallel and Distributed Computing*, vol. 74, no. 10, pp. 2899 – 2917, 2014.
- [163] Slurm Workload Manager, “Simple Linux Utility for Resource Management,” <https://slurm.schedmd.com/documentation.html>.
- [164] Juniper Research, “White paper: ‘Roam Like At Home’ impact explained,” <https://www.juniperresearch.com/document-library/white-papers/roam-like-at-home-impact-explained>.

- [165] L. M. Contreras, L. Cominardi, J. M. Pérez, and C. J. Bernardos, “Applicability of SDN and NFV Techniques for a Virtualization-Based Roaming Solution.,” *Journal of Network and Systems Management volume*, vol. 28, pp. 576 – 604, 2020.